Systems-wide analysis of isoform specific RAL biology

Thesis submitted in accordance with the requirements of the University of Liverpool for the Degree of Doctor in Philosophy

By

Yasmina Marielle Elizabeth Sahraoui

April 2021

Systems-wide analysis of isoform specific RAL biology

Yasmina Marielle Elizabeth Sahraoui

Abstract

RAL proteins (RALA and RALB) are small GTPases that operate downstream of the proto-oncogene RAS. They are known to activate a signalling network controlling membrane trafficking, cell migration and gene transcription. Whilst RAL isoforms are recognised as having overlapping but non-redundant functions, despite their high degree of structure homology, in general the isoform specific contributions of RAL signalling remain poorly understood.

Most of our understanding of RAL biology has derived from ectopic expression studies. However, the general consensus is that studying endogenous protein levels is highly advantageous. Here, the generation of a panel of gene edited colorectal cancer cell lines, harbouring different RAS mutations, enabled profiling of the wider network responses influenced by each RAL isoform as an effector of RAS.

Using RNA sequencing, the transcriptome of isoform specific RAL knockout isogenic cell lines was characterised. A subset of genes expressed in both cell lines that showed differential expression was identified. In order to infer likely protein expression differences as a result of RAL isoform loss, the RNA sequencing analysis was integrated with proteomic data obtained from RPPA technology. The combination of these two methodologies highlighted RAL responsive proteins that provide evidence the RALGEF/RAL pathway is involved in crosstalk between the other well-known RAS effector pathways, particularly proteins involved in MAPK signalling. This also extended to expression changes in proteins involved in RAS-driven processes such as cell proliferation.

Turning attention to RAL itself, activity assays revealed an increase in activity levels of the remaining RAL isoform in the knockout clones, compared to the parental, with significant RALB activity increases recorded in the RALA knockout clones across all the RAS mutant cell lines tested. It is likely that this compensatory activity results in functional redundancy for both proteins. In addition to RAL activity levels, combining targeting SRM-based proteomics with an isotope-labelled protein standard strategy enabled the most accurate measurement of cellular RAL isoform abundance to date. On the whole, this panel of colorectal cancer cell lines require sufficient levels of both RAL isoforms for proper signalling, with RALB expression being slightly more dominant than RALA. This again supports the idea of high levels of signalling compensation between the two isoforms in colorectal cancer, with any differences in RAL isoform function likely extending from different subcellular locations. Using genome-wide and protein expression/activity analysis techniques has allowed large scale profiling of isoform specific RAL signalling. Together, the data presented in this thesis provides a novel insight into RAL biology in a colorectal cancer setting.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Ambic	Ammonium Bicarbonate
ATCC	American Type Culture Collection
ATP	Adenosine Triphosphate
AUC	Area under curve
BCA	Bicinchoninic acid protein assay
Bp	Base pairs
CRC	Colorectal cancer
EGF	Epidermal growth factor
eGFP	Enhanced green fluorescent protein
FBS	Foetal bovine serum
GAP	Guanine activating protein
GDP	Guanosine diphosphate
GEF	Guanine exchange factor
gRNA	Guide ribonucleic acid
GTP	Guanosine triphosphate
GO	Gene ontology analysis
HDR	Homology directed repair
HR	Homologous recombination
HVR	Hypervariable region
IAM	Iodoacetamide
KD	Knockdown
kDa	Kilodalton
KI	Knock-in
KO	Knockout
LC	Liquid chromatography
MAPK	Mitogen-Activated Protein Kinases
Min	Minutes
mRNA	Messenger ribonucleic acid
MS/MS	Tandem mass spectrometry
m/z	Mass/charge
NHEJ	Non-homologous end joining

PAR	parental
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PSAQ	Protein standard absolute quantification
RA	RAL/RAS association domain
RBD	RAL binding domain
REM	RAS exchanger motif
RNAi	RNA interference
rpm	Revolutions per minute
RTK	Receptor tyrosine kinase
SDS-PAGE	SDS polyacrylamide electrophoresis
siRNA	Small interfering RNA
V	Volts/voltage
WT	Wild type

Acknowledgements

The past four years undertaking this PhD has been the most challenging yet rewarding time that I have experienced. Any and all achievements have required support and guidance, both of which I have been lucky to receive throughout the course of this PhD.

I would like to thank my supervisor Professor Ian Prior for not only giving me the opportunity to carry out my PhD in his lab but also for his patient guidance, invaluable insight and personal support over the years. I also want to thank him for giving me the opportunity to attend a conference hosted by The Biochemical Society at the University of Cambridge and a FASEB conference in Olean, New York, both of which I enjoyed immensely and are experiences I will remember for the rest of my life.

I am grateful for the advice and friendship of my colleagues in the department, past and present. Particularly, I would like to thank Hannah Elcocks for being the best Leo, a Libra girl could ever need. Thanks, Hannah, for listening to my rants, assisting me with my work and getting songs stuck in my head for days. I would like to pay my special regards to Dr Fiona Hood for both her professional and personal support during the course of my PhD. I would like to express my gratitude to Dr Nicholas Harper for his assistance with the CRISPR work. I also wish to thank everyone in the MCSUIP labs who have made 5th floor such a fun and welcoming working environment.

Finally, a special thanks goes to my parents, family and my partner, for their unwavering support and encouragement. In particular, my mum and nan who have sat through hours of practice presentations, spent days reading every piece of work and what has probably seemed like weeks of moaning when experiments wouldn't go to plan. Thank you to you both, as this PhD could not have been completed without you.

Chapter 1: Introduction

1.1 RAL small GTPases

1.1.1 Evolutionary Biology and Structure of RAL

To date, over 250 small GTPases have been identified as being part of the RAS superfamily (Cox and Der, 2010). Differences in the sequence, structure and function of these small GTPases allow them to be divided into five main groups, one of which is the RAS family. The RAL small GTPases are members of this RAS branch (Figure 1.1A) and due to the high degree of sequence similarity, are the closest relative of the RAS proteins, hence the basis for their names RAL (RAS-like).

The RAL gene was first identified in 1986 during a search for genes related to RAS in a simian cDNA library established from immortalised B-lymphocytes (Chardin and Tavitian, 1986). Three years later, the human RAL orthologs; RALA and RALB were identified using the simian RAL cDNA as an oligonucleotide probe in a human cDNA library (Chardin and Tavitian, 1989). Since then, RALA and RALB isoforms have been found in all vertebrate species (van Dam, Bos and Snel, 2011), as well as single RAL genes present in *C. elegans* (RAL-1) and *Drosophila* (RALA) (Frische *et al.*, 2007) (Figure 1.1A).

Despite the genes for RALA and RALB residing on different chromosomes, chromosome 7 and 2 respectively, the isoforms share >80% overall amino acid sequence identity (Figure 1.1B) (Gentry *et al.*, 2014). The tertiary structure of both proteins consists of an 11-amino acid N-terminal followed by a G-domain, containing two switch regions, and a C-terminal membrane targeting sequence (Figure 1.1C). Whilst the switch I (SI) and switch II (SII) regions of the G-domain, responsible for regulator and effector interactions, share almost 100% sequence identity, the C- terminals differ considerably, with only 50% similarity. This hypervariable region contains important post-translational phosphorylation sites that play an essential role in membrane association and localisation and consequent effector interaction.

A)

B)

Overall sequence identity				
	RALA (Hs)	**RALB (Hs)**	RAL1 (Dm)	RAL1 (Ce)
RALA (Hs)	-	82	72	65
RALB (Hs)	82	-	71	61
RAL1 (Dm)	72	71	-	67
RAL1 (Ce)	65	61	67	-

C)

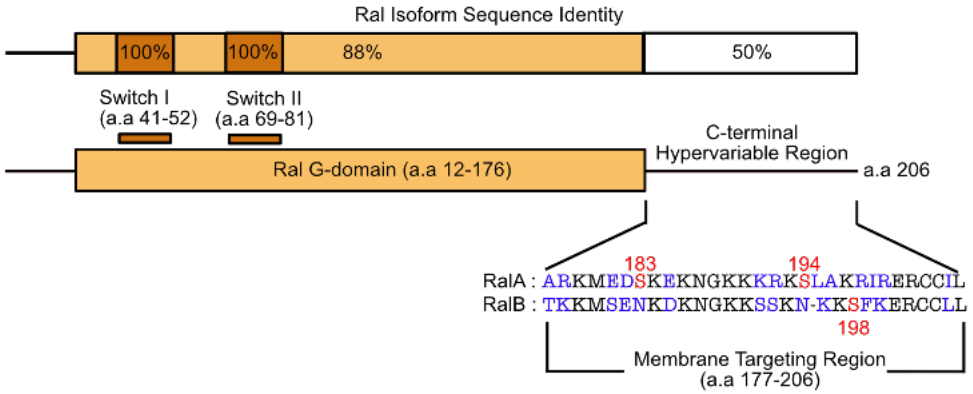


Figure 1.1 Conservation of RAL small GTPases.
A) RAL is part of the RAS branch of the RAS small GTPase superfamily. RALA and RALB isoforms are shown here alongside the four key RAS proteins and other members of the RAS family. RAL isoforms are found in all vertebrate species, whilst only a single RAL ortholog is present in *D. melanogaster* (Dm) and *C. elegans* (Ce). **B)** Human and invertebrate RAL orthologs display strong sequence identity. **C)** Schematic of RAL small GTPases and the amino acid differences between the two isoforms RALA and RALB. The G domains of the RAL isoforms share 88% sequence identity. The C-terminal hypervariable region contains key phosphorylation sites (red) that regulate isoform subcellular location and effector interaction. Thus, they exhibit only 50% sequence identity. Dendrograms and sequence identities were adapted from (Gentry *et al.*, 2014).

1.1.2 RAL activation and regulation

Like other small GTPases, RAL cycles between a GTP-bound active state and GDP-bound inactive state (Figure 1.2A) (Gentry *et al.*, 2014). As mentioned previously, the SI and SII regions of the G-domain are involved in regulator and effector recognition and the conservation of these two switch sequences can be found in the RAL proteins of *C. elegans* and *Drosophila*, supporting the interaction of RAL with conserved regulator and effector proteins.

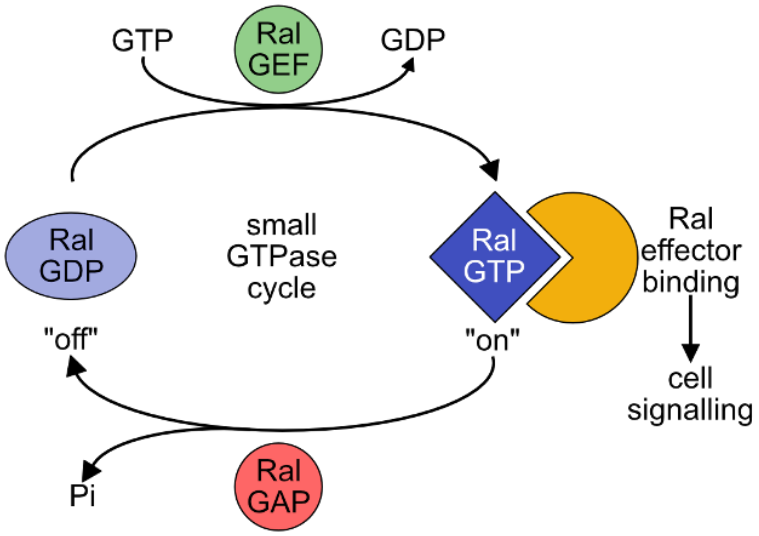
Similar to RAS, the intrinsic GTPase activity of the RAL proteins is very weak (Goody, Frech and Wittinghofer, 1991). Therefore, GDP-GTP exchange and GTP hydrolysis is accelerated by the presence of RAL guanine nucleotide exchange factors (RALGEFs) and RAL GTPase activating proteins (RALGAPs). With the intracellular levels of GTP ~10 times higher than GDP (Goody, Frech and Wittinghofer, 1991), the formation of RAL-GTP is favoured by RALGEFs, that subsequently stimulate the exchange of GDP for GTP. RALGAPs catalyse the hydrolysis of the bound GTP, returning RAL back to its inactive conformation.

RALGEFs

Currently, there are seven known RALGEFs that can be divided into two groups (Figure 1.2B). The first group includes RALGDS, RGL1, RGL2, and RGL3. These proteins rely on RAS activation by binding directly to the effector binding region of active RAS. Whereas the second group (RGL4, RALGPS1 and RALGPS2) can be activated in a RAS-independent manner (Gentry *et al.*, 2014).

RAL guanine nucleotide dissociation stimulator (RALGDS) was discovered in the early 1990s by yeast two-hybrid screening with the aim of identifying RAS effectors

A)



B)

	GENE NAME	PROTEIN NAME
RALGEFs	RALGDS	RALGDS
	RGL1	RGL
	RGL2	RGL2
	RGL3	RGL3
	RGL4	Rgr
	RALGPS1	RALGEF2
	RALGPS2	RALGPS2
RALGAPs	RALGAPA1	RALGAPalpha1/TULIP1
	RALGAPA2	RALGAPalpha2/AS250
	RALGAPB	RALGAPB

Figure 1.2 Small GTPase RAL GDP-GTP cycle.
A) Regulation of RAL is carried out by selective GEFs and GAPs that accelerate GTP hydrolysis, which promotes the formation of active GTP bound and inactive GDP bound RAL. **B)** Known regulators of the RAL GDP/GTP cycle.

(Albright *et al.*, 1993; Hofer *et al.*, 1994). RALGDS was the first RALGEF to be identified and therefore the most characterised. It has been shown to be RAL specific (Albright *et al.*, 1993), catalysing nucleotide exchange on both RALA and RALB but not on other small GTPases. Further yeast two-hybrid library screen studies identified three more RALGEF proteins named RGL1 (RALGDS-like), RGL2 and RGL3 (Kikuchi *et al.*, 1994; Isomura *et al.*, 1996; Shao and Andres, 2000; Wolthuis *et al.*, 1996). These RALGEFs were found to contain a similar domain architecture that includes a RAS exchanger motif (REM) domain at the N-terminal, followed by a CDC25 homology domain and a C-terminal RAS-association (RA) domain (Ferro and Trabalzini, 2010). The CDC25 homology domain and the catalytic domains of RASGEFs share sequence identity, allowing activated RAS to recruit RALGEFs onto the plasma membrane and deliver them to RALA and RALB (van Dam, Bos and Snel, 2011). As well as HRAS, KRAS and NRAS, other RAS family small GTPases are able to bind and activate RALGEFs through the RA domain (Rodriguez-Viciana, Sabatier and McCormick, 2004). Furthermore, a complementary mechanism of RALGEF activation, specifically RALGDS, has been shown to occur through the more familiar RAS target protein, PI3K (Tian *et al.*, 2002). The activation of PI3K promotes the binding of the PDK1 protein kinase to the N-terminus of RALGDS, which subsequently alleviates the inhibitory effect of the N-terminus on its catalytic domain allowing it to bind to RAL.

The second group of RALGEFs, RGL4, RALGPS1 and RALGPS2 also contain the CDC25 homology domain but lack the REM and RA domains, uncoupling them from association with RAS and the RAS family of small GTPases (D'Adamo *et al.*, 1997; de Bruyn *et al.*, 2000; Rebhun, Chen and Quilliam, 2000). Instead, RALGPS1 and RALGPS2 contain a C-terminal pleckstrin homology (PH) domain that has been shown to be involved in membrane targeting and RAL activation (Rebhun, Chen and Quilliam, 2000). Whilst RGL4 does contain this CDC25 homology domain, a well-defined RA or PH domain is absent and its activation mechanism is still poorly understood. Although the RALGEFs mentioned previously are highly selective RAL activators, RGL4 is known to activate other RAS family small GTPases (Osei-Sarfo *et al.*, 2011).

RALGAPs

Like RALGEFs, RALGAP activity was also reported in the early 1990s (Emkey, Freedman and Feig, 1991). However, molecular identification of these proteins has only recently been achieved. Using a GTPase-deficient, constitutively active RALA mutant for affinity chromatography, two RALGAP complexes, RALGAP1 and RALGAP2 were identified from brain cytosol (Shirakawa *et al.*, 2009). RALGAPs (Figure 1.2B) are heterodimers each consisting of a RALGAPβ regulatory subunit and one of two RALGAPα1 and α2 catalytic subunits (Chen et al. 2011) (Chen *et al.*, 2011b). As with the RALGEFs, RALGAPs are RAL specific, only accelerating the GTPase activity of both RALA and RALB but no other small GTPases (Shirakawa *et al.*, 2009; Chen *et al.*, 2011b; Martin *et al.*, 2014). Relatively little is known about these RAL specific GAP proteins. However, the subunit RALGAPβ was found to play an important role in RAL-mediated spindle formation regulation during mitosis (Personnic *et al.*, 2014) and RALGAPα2 (previously known as AS250; AKT substrate of 250 kDa) is known to undergo phosphorylation by AKT, reducing its interaction with RALA (Chen et al. 2011).

Post-translational modification

RALA and RALB exhibit the most significant sequence divergence in their C-terminal hypervariable region (Figure 1.1C). Like RAS, RALA and RALB terminate in a CAAX motif sequence (C = cysteine, A = aliphatic amino acid, X= any amino acid), which undergoes a series of posttranslational modifications that are necessary for membrane association (Gentry *et al.*, 2015). The terminal X residue at the end of the CAAX motif, determines protein prenyltransferase specificity. The CAAX motif are CCIL and CCLL for RALA and RALB respectively. In the case of both RAL isoforms, the presence of the leucine (L) residue signals for the addition of an isoprenoid lipid to the first cysteine (C), which is carried out by geranylgeranyltransferase type I (GGTase-I). Modifications of the CAAX motif are critical for normal RAL function and mutation of the cysteine residue preventing these necessary changes, disrupts RAL association with the plasma membrane (Falsetti *et al.*, 2007). Whilst both isoforms can be found at the plasma membrane and

endomembranes of the cell, this sequence divergence results in distinct subcellular localisation that subsequently contributes to functional difference between the two proteins.

In addition to the prenylation of the CAAX motif, RAL subcellular location is also regulated by reversible post-translational modifications through phosphorylation of RAL at the C-terminal region with distinct protein kinases. Aurora-A kinases and protein kinase A (PKA) were found to phosphorylate RALA at residue S194 (Wu *et al.*, 2005; Lim *et al.*, 2010), with dephosphorylation at the same residue carried out by protein phosphatase 2A (Sablina *et al.*, 2007). Phosphorylation of RALA at S194 by Aurora A alters subcellular location of RALA from the plasma membrane to internal membranes, allowing for enhanced interaction with the effector RALBP1 (Lim *et al.*, 2010). Aurora A phosphorylation of RALA also promotes translocation to the mitochondrial outer membrane, where it again recruits RALBP1 to stimulate mitochondrial fission (Kashatus *et al.*, 2011).

Studies have revealed RALB is similarly regulated through phosphorylation of residue S198 by protein kinase C alpha (PKCα), relocating it from the plasma membrane to the endocytic vesicles (Martin *et al.*, 2012; Wang *et al.*, 2010). In addition to changes in subcellular location, phosphorylation of both isoforms is associated with a change in effector utilisation. Unphosphorylated RALB preferentially binds to the effector Sec5, whilst S198 phosphorylation leads to preferential association with RALBP1 (Gentry *et al.*, 2014). Together, alteration of RALs GTP and GDP bound state by GEFs and GAPs, C-terminal posttranslational modifications and phosphorylation provide spatial and temporal regulation of RAL and allow cells to respond to different stimuli in unique ways.

1.2 RAL function

1.2.1 RAL effectors

Following activation, GTP bound RAL carries out various biological functions, such as cytoskeleton organisation, membrane trafficking, transcription regulation and kinase signalling (Figure 1.3), via different effector binding proteins. Since RALA and RALB are regulated by the same set of proteins and have the same enzymology, they are able to interact with the same set of effectors. The best characterised proteins being RALBP1/RLIP76 and the Sec5/Exoc2 and Exo84/Exoc8 subunits of the octameric exocyst complex. As described earlier, a combination of regulatory processes leads to differences in subcellular localisation. Therefore, even though the effector proteins are the same, distinct functions can arise.

RALBP1/RLIP76

RAL binding protein 1 (RALBP1) was the first effector of RAL to be discovered in screens searching for proteins that preferentially bound to RALA (Cantor, Urano and Feig, 1995; Jullien-Flores *et al.*, 1995; Park and Weinberg, 1995). The most extensively studied effector of RAL, RALBP1 is a 76 kDa protein that includes a binding domain that is able to bind to both switch regions of GTP-RAL (Fenwick *et al.*, 2010). RALBP1 itself binds to a diverse group of proteins and bears RHOGAP activity for CDC42 and RAC1 (members of the RHO branch of the RAS superfamily) small GTPases (Wennerberg, Rossman and Der, 2005). These proteins are involved in cytoskeletal changes such as actin dynamics, filopodia and lamellipodia formation and membrane ruffling (Park and Weinberg, 1995). Overexpression of RALBP1 has also been found in different human cancers, with RALBP1 suppression impairing tumourigenic growth *in vivo* (Singhal *et al.*, 2007), regulating oral cancer cells via AKT (Ieong, Ma and Lai, 2019) and inducing resistance to apoptosis in breast and prostate cancer (Singhal *et al.*, 2019; Yang *et al.*, 2015). However, it should be noted that any phenotypes attributed to RALBP1, are not necessarily linked with RAL signalling.

RALBP1 has additional functions that regulate endocytosis and signal transduction. The N-terminus of RALBP1 can associate with the AP2 adaptor

complex, a regulator of clathrin-mediated endocytosis (Jullien-Flores *et al.*, 2000), whilst the C-terminal region interacts with Reps1 and Reps2 (Yamaguchi *et al.*, 1997; Ikeda *et al.*, 1998; Nakashima *et al.*, 1999). These are important proteins involved in receptor tyrosine kinase-regulated endocytosis (Cullis *et al.*, 2002; Morinaka *et al.*, 1999). Prevention of endocytosis during mitosis has also been linked with the RALA induced activation of RALBP1 (Rossé *et al.*, 2003). The C-terminus of RALBP1 associates with cyclin B1 and subsequently form complexes with CDK1, which results in CDK1 phosphorylation of the endocytic protein Epsin.

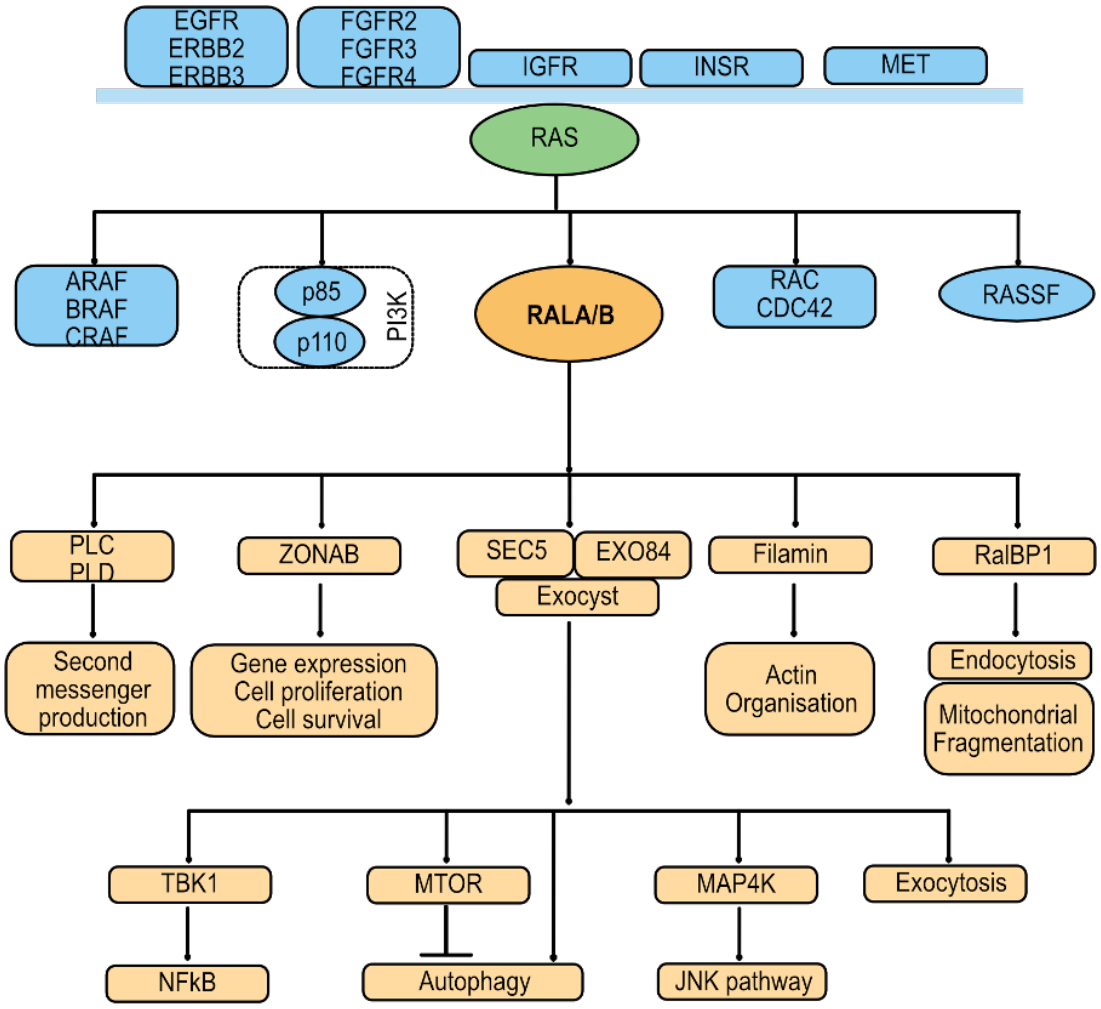


Figure 1.3 RAL and its effectors and effector functions.
RAS activated RAL can regulate a variety of downstream effectors that modulate numerous cellular processes. Cell cycle progression is regulated through PLD cytokinesis, and gene transcription through ZONAB. RAL associates with the subunits Sec5 and Exo84 to regulate exocyst processes while cytoskeletal changes are modulated by Filamin. RALBP1 acts as a scaffold for proteins involved in endocytosis.

Several other RAL-driven processes require RALBP1 as a key effector. Many of these studies utilise RAL isoforms that have selective mutations impairing effector interaction to determine protein function. For example, the D49N substitution impairs RAL-RALBP1 but not RAL-Sec5/Exo84 binding, whereas a D49E mutation results in the opposite effect (Cantor, Urano and Feig, 1995; Moskalenko *et al.*, 2002; Moskalenko *et al.*, 2003). Consequently a D49E mutation was important in discovering that RALBP1 was a critical effector for the RALB-specific function of invadopodia formation, after rescuing the loss of endogenous RALB following shRNA silencing (Neel *et al.*, 2012).

RALA inhibits TGFβ-mediated growth arrest in epithelial cell through the mislocalisation of the cyclin-dependent kinase inhibitor p27 KIP1 (Tazat *et al.*, 2013). This process appears to rely on the RHOGAP domain of RALBP1.

Recruitment of RALBP1 by RALA to the mitochondria is required for mitochondrial fission during mitosis (Kashatus *et al.*, 2011). Fission is necessary for the proper distribution of mitochondria to daughter cells during cell division and is controlled by the GTPase DRP1 present on the outer mitochondrial membrane. Once recruited, RALBP1 acts as a scaffold that facilitates the phosphorylation of DRP1 by cyclin B/CDK1 complexes. Loss of mitochondrial fission was observed following suppression of either RALA or RALBP1.

The Sec5 and Exo84 subunits of the exocyst

Both Sec5(Exoc2) and Exo84(Exoc8) are key subunits of the exocyst complex, which plays an essential role in the generation of exosomes and cell migration. This multiprotein complex is responsible for trafficking various secretory vesicles to specific regions of the plasma membrane (He and Guo, 2009).

The association of these two subunits with RAL has been found to be important in exocytosis (Moskalenko *et al.*, 2002; Moskalenko *et al.*, 2003; Sugihara *et al.*, 2002; Camonis and White, 2005). Sec5 binds to the switch I region of RAL (Fukai *et al.*, 2003) and mediates the delivery of membrane proteins in epithelial cells as well as vesicle secretion in neuronal cells following direct binding to RALA (Moskalenko *et al.*, 2002). The Exo84 subunit

binds to both switch regions of activated RAL (Jin *et al.*, 2005) and is essential for RAL-mediated assembly of the hetero-octameric exocyst complex (Moskalenko *et al.*, 2003). The subcellular localisation of the exocyst is regulated by RAL promoting Sec5-paxillin interactions and the subsequent assembly of the whole exocyst complex including Exo84 (Spiczka and Yeaman, 2008).

RAL is also capable of interaction with Sec5 and Exo84 to regulate cellular processes that are independent of their roles in exocytosis. The employment of Sec5 by RALB is important in activating the innate immune response (Chien *et al.*, 2006). The interaction between RALB and Sec5 results in Sec5 binding with the protein kinase TBK1, which is known to regulate NF-κB signalling.

On the other hand, during nutrient starvation or the presence double-stranded DNA, RALB engages with the EXO84 subunit and the exocyst, which in turn upregulates autophagosome formation (Shi *et al.*, 2012; Bodemann *et al.*, 2011). The interaction between RALB and EXO84 triggers the assembly of the serine/threonine kinase ULK1 and the Beclin1-Vps34 autophagy initiation complex. This differential interaction between RALB and either Sec5 or Exo84 is achieved molecularly by ubiquitylation of RALB at Lysine 47, which is controlled by the deubiquitylase USP33 (Simicek *et al.*, 2013). Here, ubiquitination inhibits EXO84 binding and facilitates SEC5 interaction. These data suggest that RALB in association with different exocyst subunits acts as a regulatory hub that trigger specific cellular responses supporting systemic pathogen recognition and removal.

Transcription Factors and other effectors

RAL also has other lesser-characterised effectors, one of which is phospholipase D1 (PLD1) (Luo *et al.*, 1998; Kim *et al.*, 1998). Unlike other effectors, the association of PLD1 with RAL is not GTP dependent. Instead of the effector binding domain, the association is with the N-terminus of RAL. PLD1 is an important second messenger protein involved in the conversion of phosphotidylcholine to phosphatidic acid and choline following G-protein coupled receptor (GPCR) stimulation. RALA-PLD1 binding is also important in the localisation of p27; a cyclin-dependent kinase inhibitor, thus promoting

proper TGFβ signalling (Tazat *et al.*, 2013). Additionally, RALA is necessary for PLD1-mediated activation of mTORC1 signalling (Xu *et al.*, 2011).

Filamin is an important protein of the cytoskeleton, playing a role in actin crosslinking and the formation of lamellipodia. RALA was found to interact with filamin and induce filopodia formation in Swiss-3T3 cells (Ohta *et al.*, 1999).

Lastly, RAL has been shown to engage with several transcription factors. At high cell densities, RALA engages with ZONAB (zonula occludens 1-associated nucleic acid binding protein), unlocking ZONAB targets for transcription in MDCK cells (Frankel *et al.*, 2005).

Table 1.1 RAL effectors and functions.

RAL ISOFORM	EFFECTOR/ PROCESS AFFECTED	METHOD/technique used
RALA	RALBP1 - endocytosis	RNAi
	RALBP1 - TGFβ-mediated growth	RNAi- shRNA ectopic expression
	RALBP1 – mitochondrial fission	RNAi- shRNA ectopic expression
	PLD1 – TGFβ signalling	RNAi- shRNA ectopic expression
	PLD1 – mTORC signalling	RNAi – siRNA
	Filamin - filopodia	ectopic expression
RALB	RALBP1 – invadapodia	RNAi- shRNA
	SEC5 – innate immune response	RNAi – siRNA
	EXO84 – autophagosome formation	RNAi – siRNA
RALA and RALB	SEC5 + EXO84 - Exocyst	ectopic expression
	ZONAB	ectopic expression

Other examples of transcription factor regulation by RAL, include the phosphorylation of c-JUN through JNK (de Ruiter *et al.*, 2000), ROS-induced, JNK-dependent regulation of FOXO (forkhead box O) transcription factors by promotion of a JIP1 scaffold complex (van den Berg *et al.*, 2016), STAT3 activation following EGF stimulation through the non-receptor tyrosine kinase SRC (Goi *et al.*, 2000) and B cell receptor (BCR)-controlled activation of JUN/ATF2 and NFAT (de Gorter *et al.*, 2007).

Although RALA and RALB are structurally similar and are known to bind to the same set of effector proteins, they have been observed as having isoform-specific functions (Table 1.1). Nonetheless, it seems that both RAL proteins are involved in a wide variety of biological processes. Whilst these studies have greatly assisted in determining the role of RAL under normal cellular physiology, it should be noted that all interactions and observations of RAL were determined following the use of gene knockdown and gene overexpression techniques. Although the use of RNAi, is a direct approach to gene disruption and loss of function, off-target effects are a growing concern. The siRNA induced silencing of non-target mRNA transcripts is particularly an issue (Carthew and Sontheimer, 2009), as potentially hundreds of transcripts (Sigoillot and King, 2011) can be repressed resulting in phenotypes that dominate over the intended target phenotype (Franceschini *et al.*, 2014). Overexpression of proteins poses its own problems and limitations, such as the burden protein production puts on the cellular resources, stoichiometric imbalance, promiscuous protein interaction and pathway modulation (Moriya, 2015). All of these may potentially hinder the fitness of the cell and cause cellular toxicity.

1.2.2 RAL and cancer

Since RALGEFs act as direct downstream effectors of activated RAS, it was suggested that RAL signalling may contribute to RAS-driven tumourigenesis. Perhaps surprisingly, early studies in NIH 3T3 mouse fibroblasts found that only the RAF effector pathway and not RAL was sufficient to mediate RAS-driven cellular transformation (Urano, Emkey and Feig, 1996). However, later studies exploring the role of RAL in RAS-dependent transformation in immortalised human cell supported a more significant role for RAL GTPases as effectors of RAS in human cancers (Hamad *et al.*, 2002). Following depletion of either RAL isoform via RNAi, it was further demonstrated that RALA but not RALB is necessary for anchorage-independent growth of cancer cells whereas RALB is required for tumour cell survival (Chien and White, 2003).

It should be noted that since RALGEFs can be activated and regulated by RAS independent mechanism, any involvement of RAL GTPases in cancers, where RAS mutations are not present is not surprising. Indeed, a critical role for RAL in both tumourigenesis and metastasis in many types of human cancer has been established.

Pancreatic ductal adenocarcinoma (PDAC)

Pancreatic cancer is one of the most popular models for studying the role that RAL plays in human cancer, due to the high frequency of activating KRAS mutations present (>90%, COSMIC database) and the activation of RAL seen in both tissue samples and tumour cell lines (Jones *et al.*, 2008; Lim *et al.*, 2005; Lim *et al.*, 2006). Within these studies, activation of RAL was found to be more common than that of either ERK or AKT, suggesting a key role for RAL signalling downstream of oncogenic KRAS besides the RAF and PI3K effector pathways.

Stable shRNA suppression of RALA, but not RALB, in established primary tumours resulted in tumour regression, indicating persistent RALA signalling is necessary for PDAC tumour maintenance in addition to tumour initiation (Lim *et al.*, 2006). Distinct roles of RAL in RAS-driven cancers has been confirmed by another study showing active KRAS signalling to RALB but not

RALA is involved in invadopodia formation in human pancreatic cancer cells (Neel *et al.*, 2012). This supports the idea of a role for RALB in metastasis, since invadopodia are known to be involved in the secretion of matrix metalloproteases (MMPs) during tumour cell invasion.

Other members of the RAL pathway have been found to play a role in PDAC. The RALGEF protein RGL2 is overexpressed in patient tumours and has been shown to be a requisite for anchorage independent growth and invasion of PDAC tumour cells (Vigil *et al.*, 2010). Indeed, RNAi-mediated depletion of RGL2 led to a significant decrease in both RAL isoform activation. It should be highlighted that involvement of RAL regulators in cancer progression does not necessarily require RAL itself and interestingly, expression of constitutively active RALA did not rescue soft agar growth after the loss of RGL2. This may be due to RGL2 having non-RAL dependent regulatory functions or that the interaction between RALA and RGL2 is critical for the regulation of anchorage-independent growth that was observed (Vigil *et al.*, 2010). The idea of different roles carried out by RAL isoforms in cancer is supported by the co-localisation of RGL2 with RALB but not RALA at the leading edge of CFPac-1 PDAC cells during migration. Again, loss of RGL2 is accompanied by loss of RALB from the leading edge, suggesting that PDAC migration and invasion may rely on RGL2/RALB signalling.

As well as consistent hyperactivity of RALA and RALB seen across pancreatic cell lines and tumour samples, loss of RAL control was found to be a risk factor for pancreatic cancer development *in vivo* (Beel *et al.*, 2020). Here, deregulation of the GTPases κB-RAS1 and κB-RAS2 (NF- κB inhibitor interacting RAS-like protein 1/2) results in upregulation of RAL activity through the loss of RALGAP function, promoting acinar-to-ductal metaplasia (ADM) during tumour initiation and progression.

Colorectal carcinoma

As in pancreatic cancer, a high frequency of oncogenic KRAS mutations are present in colorectal cancer (CRC) tumours (45%, COSMIC database). Proliferation of KRAS mutated colorectal cancer cell lines was found to be significantly reduced following RALA knockdown (Győrffy *et al.*, 2015).

Furthermore, identification of a RALA-responsive gene expression signature from these cell lines, correlated with progression-free survival of colorectal cancer patients in five independent data sets.

RALA and RALB have been found to be upregulated in both colorectal cancer cell lines and patient samples (Martin *et al.*, 2011). Knockdown of RALA resulted in a decrease in soft agar colony growth, whereas RALB knockdown had the opposite effect and led to enhancement of anchorage-independent growth of cancer cells. This contrasting phenotype was found to be attributed to the differential use of common RAL effector proteins. Both RAL isoforms required interaction with RALBP1 but with different subunits of the exocyst. RALA binds to Exo84, whereas RALB interacts with Sec5 to carry out their regulation of anchorage-independent growth in colorectal cancer (Martin *et al.*, 2011).

Lung adenocarcinoma

In contrast to colorectal and pancreatic cancer, RALA and RALB have been found to have redundant effects on the growth of non-small cell lung cancer (NSCLC) tumour cells. Depletion of both RALA and RALB or either one alone, reduced anchorage independent growth in KRAS mutant NSCLC cell lines (Ihle *et al.*, 2012). Analysis of RAL expression in NSCLC patient tumours found that both RAL isoforms had higher expression in those with a KRAS mutation and was associated with poor survival (Guin *et al.*, 2013). However, only RALA activity levels were found to be elevated in KRAS-mutant NSCLC cell lines, suggesting it's RALA that plays a role in tumour growth. Interestingly, a relationship between KRAS mutation and RAL dependence has been established across many NSCLC cells, where KRAS G12C or G12V mutations have been found to carry an increased sensitivity to RAL depletion, supporting the importance of RAL signalling for tumour growth downstream of RAS (Ihle *et al.*, 2012).

Other cancers

Increased activation of RALA and RALB has been detected in human bladder cancer cell lines (Smith *et al.*, 2007; Saito *et al.*, 2013), with similar tumour

growth functions but opposite roles in motility found in the KRAS mutant UMUC3 line (Oxford *et al.*, 2005). This motility promoting function of RALB was found to be dependent on its post-translational phosphorylation modification by protein kinase C (Wang *et al.*, 2010).

RALA was recorded as being significantly overactivated in hepatocellular carcinoma (HCC) cells and tissues, with suppression of expression leading to a decrease in viability and invasiveness of HCC cells (Ezzeldin *et al.*, 2014).

Analysis of RAL activation in human melanoma cells showed a consistent increase in total and activated levels of RALA and were shown to be necessary for the tumourigenic growth of melanomas (Zipfel *et al.*, 2010). Ectopic expression of the RALGEF RGL2, mutated to mimic RAS activation, sufficiently promoted anchorage independent growth of melanocytes, whilst the expression of a RALB mutant, that blocks RALGEF function, impaired the growth of transformed melanocytes (Mishra *et al.*, 2010).

Elevated RAL expression has also been observed in malignant peripheral nerve sheath tumours (Bodempudi *et al.*, 2009), prostate cancer (Yin *et al.*, 2007) and medulloblastomas (Ginn *et al.*, 2016). Squamous cell carcinoma (SCC) is one exception, where RALA was found to suppress rather than promote tumour progression in SCC cell lines (Sowalsky *et al.*, 2010) and a RALA gene expression signature supporting this tumour suppressor function was downregulated in cancer tissues (Smith *et al.*, 2012).

Together, these studies provide increasing evidence of RAL GTPases playing a key role in cancer cell growth. Across most cancer types examined, the expression and activation levels of RALA and RALB was increased in patient tumour samples compared with normal tissue. Studies utilising human cancer cell lines have established that, in general, RALA is required for anchorage-independent growth, whereas RALB is important in tumour survival and invasion. Consequently, the idea that RAL may contribute to oncogenic RAS function should not be discounted and when it comes to identifying pharmacologic approaches for correcting aberrant RAL function, both isoforms should be targeted simultaneously.

1.3 Targeting RAL signalling

1.3.1 Efforts to target RAL directly

When it comes to targeting RAL, like RAS, there are several issues that hinder the development of compounds that directly bind to and inhibit RAL activity. First, the high concentration of GTP/GDP in cells, along with the high affinity of RAL for guanine nucleotides make the development of a molecule that can compete with GTP/GDP binding extremely difficult (Cox *et al.*, 2014). Second, the protein-protein interactions in the RAL signalling cascade are harder to target in comparison to the enzymatic activity of protein kinases (Arkin, Tang and Wells, 2014). Lastly, activated RAL (GTP-RAL) has a very "smooth" surface with no deep hydrophobic pockets suitable for tight binding of small molecules (Yan and Theodorescu, 2018). Fortunately, unlike RAS, the presence of RAL mutations in cancer are extremely rare, which allows more options for targeting (Ostrem *et al.*, 2013). Furthermore, advances in structure-based drug discovery have helped shed light on small GTPases such as RAS and RAL, allowing for development of inhibitors.

Structure-based drug discovery is based on understanding the three-dimensional structure of the target protein, which is usually determined through either X-ray crystallography or nuclear magnetic resonance (NMR), both of which have drawbacks when it comes to structure determination (Chen *et al.*, 2012). For instance, X-ray crystallography traps the target molecule in a single low-energy confirmation, which might not be physiologically relevant, whilst NMR structure determination is restricted to small proteins usually less than 30 kDa (Gauto *et al.*, 2019). Additionally, the idea that small GTPases exist solely in either an inactive GDP-bound or active GTP-bound state was disproved when NMR studies of HRAS uncovered the existence of at least two conformations of the active form (Shima *et al.*, 2010; Shima *et al.*, 2013). The concurrence of a more dynamic conformational status of small GTPases has led to several breakthroughs in developing small molecule compounds, such as the two KRAS inhibitors DACI and VU0460009 (Maurer *et al.*, 2012; Sun *et al.*, 2012). Fragment-based drug design, used to develop drugs such as these two KRAS inhibitors, is based on identifying chemical fragments that are capable of binding to small pockets on the target protein and then growing or

linking them together to produce a lead compound with a higher affinity. This approach is particularly promising for targeting small GTPases such as RAL (Zhao *et al.*, 2017).

Another breakthrough in structure-based drug design is the idea of a “stapled peptide”. This approach is based on mimicking the α-helices present in protein binding domains and applying a constrained in the form of a covalent linkage or “staple” (Walensky and Bird, 2014). The introduction of a synthetic brace (staple) aids in locking a small peptide into a specific conformation. This helps increase stability, cell permeability and target affinity of the peptide.

This technique has been applied to target RAL, in which the binding domain of the RAL effector RALBP1 was used as a template (Thomas *et al.*, 2016). The designed peptides can bind selectively to the switch I and II regions of GTP-RALB, blocking downstream effector signalling of RAL. Recent efforts to improve binding of these stapled peptides to RAL proteins has been achieved through affinity maturation of the RALBP1 binding domain (Hurd *et al.*, 2020). Several sequence substitutions were identified and used to design second-generation stapled peptides that exhibited improved selectivity for RAL GTPases. Although these stapled peptides are able to successfully compete with multiple RAL effectors, because these peptides do not bind to the inactive GDP-bound RAL, they are unable to prevent RALGEF binding and are therefore unable to inhibit RAL activation.

Another structure-based approach has led to the development of allosteric inhibitors capable of binding to inactive GDP-RAL. Using a virtual chemical library screen, small ligands that bound to an allosteric binding site on the surface of RAL were identified. Since these ligands lock RAL in the inactive GDP-bound form, they are able to break up the GTP-GDP cycle and block RAL signalling. In fact, the first of these developed RAL compounds inhibited the growth of human NSCLC cell lines both *in vitro* and *in vivo* (Yan *et al.*, 2014).

Finally, in an effort to target RAL GTPases directly, the formation of a covalent bond at a tyrosine residue (Tyr-82), inhibits RALGEF RGL2-mediated nucleotide exchange of RAL (Bum-Erdene *et al.*, 2020). Although these covalent inhibitors are not considered suitable for therapeutic drug use, they

serve as valuable probes to uncover possible binding sites for future drug development.

1.3.2 Targeting RAL localisation

In order for RAL proteins to carry out their biological functions, they must be correctly localised to the inner surface of the plasma membrane. As discussed earlier, both RAL proteins have a CAAX motif C-terminal sequence that requires modification to allow for appropriate localisation. When it comes to posttranslational modifications of RAL, unlike RAS which primarily undergoes prenylation by farnesyltransferase (FTase), RAL proteins can act as substrates for geranylgeranyltransferase (GGTase I) (Kinsella, Erdman and Maltese, 1991) and will not be subjected to prenylation following GGTI inhibitor treatment (Falsetti *et al.*, 2007). The inhibition of GGTase I and its effect on RAL signalling has been evaluated in the human pancreatic cancer cell line Mia PaCa-2 (Falsetti *et al.*, 2007). Both RAL isoforms were involved in the inhibition of anchorage-independent growth as well as cell cycle progression in these cells. The growth inhibition effect of GGTIs within RALB signalling, has also been confirmed in squamous cell carcinoma (Hamada *et al.*, 2011) and other cancer types such as glioma (Song *et al.*, 2015). A potential problem with targeting RAL through GGTase inhibition is that many other RAS family small GTPases are also substrates for GGTase I. This lack of specificity could result in toxicity in normal tissues. Fortunately, the inhibitor GGTI-2418 used in a phase I clinical trial (Ullah, Mansha and Casey, 2016) seemed to be well tolerated and the use of GGTIs for targeting RAL may still be a feasible approach.

1.3.3 Targeting the RALGEF/RAL pathway

Instead of targeting RAL itself, another logical approach would be to develop inhibitors directed at other members of the RAL pathway involved in regulation. The majority of RALGEFs mediate the activation of RAL in a RAS-dependent manner. Therefore, targeting RALGEFs should be as effective as inhibiting RAL directly, as long as the RAL signalling is a result of RAS activation. Unfortunately, targeting any GEF protein of the small GTPases is a challenge as it requires targeting protein-protein interactions. Whilst inhibitors

have been developed that target GEFs of the small GTPases, ARF, RHO and RAC1 (Schmidt *et al.*, 2002; Renault, Guibert and Cherfils, 2003; Gao *et al.*, 2004), structural information for RALGEFs is limited and making it difficult for them to be used as feasible therapeutic targets.

Another option to interfere with RAL signalling is to target downstream effector proteins that directly bind to active RAL. Currently, specific inhibitors for most RAL effectors are not available but targeting these downstream protein interactions is a sound approach that may aid in the development of novel inhibitors of RAL signalling.

1.4 Genome Editing

1.4.1 Targeted genome editing development

In order to study RAL function, generating appropriate gene "knockout" or "knock-in" cell lines through genome editing is highly desirable. Through the loss/gain of function of a target gene, protein interactors and signalling pathways can be identified and studied. Conventional gene targeting was initially achieved through the addition of a donor DNA template containing the gene of interest or mutation that would be incorporated at the target site through homologous recombination (HR) (Capecchi, 1989). Whilst HR has been used to successfully generate knockout murine embryonic stem cells (Capecchi, 2005), the rate at which it occurs is low and lower still in other cell types, making it an inefficient process to utilise exclusively.

Over recent years, there has been extensive advancements in genome editing technologies that enable precise modification of a specific locus within the genome through the introduction of a double-stranded break (DSB). The idea that DSBs could boost the efficiency of HR was first shown in yeast (Haber *et al.*, 1988) and at later dates in mammalian cells (Rouet, Smih and Jasin, 1994). Furthermore, the discovery that DNA cleavage triggers the cell's natural repair mechanism, known as non-homologous end joining (NHEJ), a donor DNA template is no longer required to disrupt the gene of interest and produce a "knockout" cell line (Wake *et al.*, 1984).

The introduction of a break in the DNA is carried out by the targeted delivery of a nuclease, with zinc finger nucleases (ZFNs) being the first to be used in

this manner (Kim, Cha and Chandrasegaran, 1996). ZFNs are endonucleases that consist of a DNA-binding domain of a zinc finger protein transcription factor, fused to a separate DNA-cleavage domain (Carroll, 2011). Here, DNA binding specificity can be provided through the customisation of the zinc finger protein to target any user-defined sequence. Transcription activator-like effector nucleases (TALENs) were the next nuclease to be developed (Miller *et al.*, 2011). Like ZFNs, TALENs contain an endonuclease domain, this time linked to two DNA-binding motifs instead of one (Joung and Sander, 2013). Whilst there are structural similarities, TALENs have an advantage over other nucleases in that the DNA recognition code can be easily engineered to target almost any sequence (Boch *et al.*, 2009). However, the substantial size of TALENs presents a challenge when it comes to packaging and delivery into cells and is especially problematic in applications that must be delivered in viral vectors (Gupta and Musunuru, 2014). Finally, the current gold-standard of genome editing, and precise modification is the CRISPR/Cas system.

1.4.2 CRISPR/Cas system

Clustered regulatory interspaced palindromic repeats (CRISPR) and the CRISPR-associated (Cas) nucleases were initially discovered as part of the bacterial adaptive immune system (Fineran and Charpentier, 2012). Any invading foreign genetic material detected by the bacteria is catalogued and incorporated into their own genome as “protospacers” and expressed as short guide RNAs (gRNAs). The CRISPR/Cas system uses a combination of these short RNAs to guide Cas nucleases to target and cleave specific foreign DNA (Barrangou *et al.*, 2007).

Various CRISPR/Cas systems are available for use as a gene editing tool but the most comprehensively studied is the type II system from the bacterial species *Streptococcus pyogenes* (Cong *et al.*, 2013; Mali *et al.*, 2013; Jinek *et al.*, 2012; Cho *et al.*, 2013). The CRISPR gene loci, present in both bacterial plasmid and chromosomal DNA, is comprised of gRNA(s) along with a Cas9 nuclease (Mojica *et al.*, 2005). This CRISPR/Cas system has been harnessed to facilitate human genome editing through human optimisation of the Cas9 nuclease, in addition to design of the relevant gRNA components. This includes fusing a trans-activating CRISPR RNA to the 20 base pair sequence

gRNA through a linker loop region to form the final single gRNA (Kim and Kim, 2014) (Figure 1.4).

Directly upstream of the Cas9 nuclease binding site on the target DNA strand is a small base pair sequence known as the protospacer adjacent motif (PAM). Different CRISPR/Cas9 systems have discrete PAM requirements, with the PAM sequence in this particular system being a 5'-NGG sequence, where 'N' is any random nucleotide (Jiang *et al.*, 2013). The gRNA recognises and hybridises to the matching part of the genome, leading to a Cas9-induced DSB three base pairs downstream of the PAM. Importantly, cleavage will only occur if a PAM sequence is present (Ding *et al.*, 2016).

As previously mentioned, subsequent repair of these DNA breaks occurs through the cellular process NHEJ (Figure 1.4). This process is highly error prone and can be exploited to introduce insertion/deletion (indel) mutations, potentially knocking out the gene, leading to premature truncation of the protein (Gupta and Musunuru, 2014). Alternatively, homology directed repair (HDR) can be utilised with the addition of a repair template, to insert a specific mutation, allowing for precise repair and the option of knock-in gene editing (Chapman, Taylor and Boulton, 2012).

1.4.3 CRISPR and RAL

CRISPR can also be used to conduct large-scale genetic loss of function experiments (Poirier, 2017). This technique known as CRISPR screening, facilitates the discovery of key genes that elicit a specific phenotype for a cell type, such as drug resistance/sensitivity (Shalem *et al.*, 2014), susceptibility to environmental toxins (Koike-Yusa *et al.*, 2014) as well as leading to a particular disease state (Chen *et al.*, 2015). The basic idea of a CRISPR screen uses precise DSBs to knockout every gene but only one gene per cell, so the resulting population has a different gene knocked out. Following a period of time in which some cells will die, and some will become the predominant cell type, next-generation sequencing is carried out to determine individual gene level changes. This allows the identification of genetic sequences that are crucial for survival under normal or specific conditions.

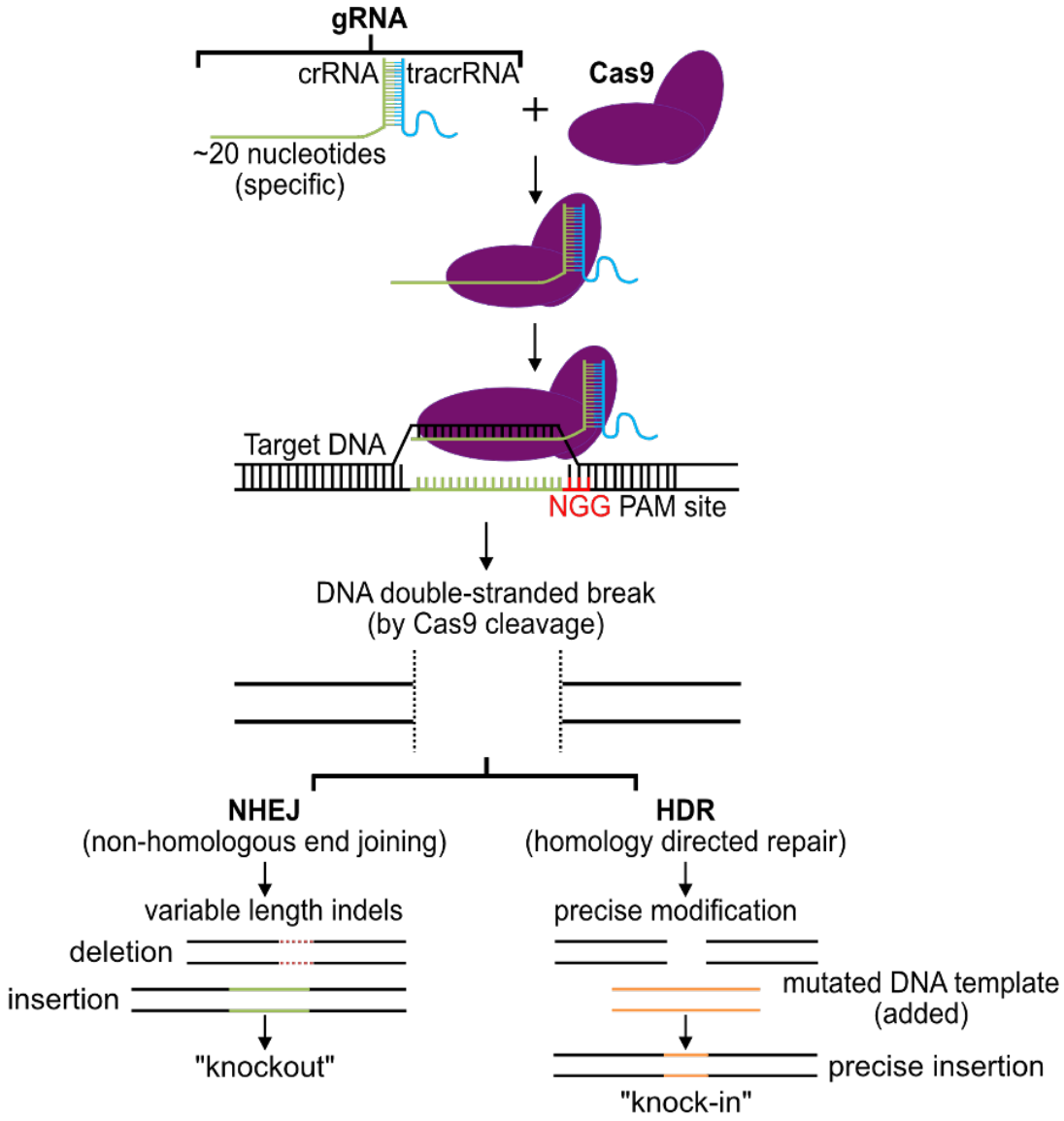


Figure 1.4 The CRISPR/Cas9 system.
Schematic of the CRISPR/Cas9 system including gRNA with 20 base pair target sequence and Cas9 nuclease that generates double strand breaks in the DNA. The two methods of repair are also depicted, NHEJ and HDR. crRNA = CRISPR RNA, tracrRNA = trans-activating CRISPR RNA, NNG = N – random nucleotide, G – guanine, PAM = protospacer adjacent motif.

Out of 1007 and 1001 CRISPR screens conducted in human cell lines, 13 and 7 of them identified RALA and RALB as hits respectively (BioGRID ORCS). One screen focusing on a peptide accumulation phenotype in an embryonic kidney cell line found that RALA may act as a suppressor of ER stress-induced apoptosis (Panganiban *et al.*, 2019). Eleven screens, focused on changes in cell proliferation, with RALA recorded as being a hit in a range of cell types including large B-cell lymphoma, neuroblastoma and endometrial cancer (Behan *et al.*, 2019), neural stem cells and glioblastoma (MacLeod *et al.*, 2019), B-cell non-Hodgkin Lymphoma (Reddy *et al.*, 2017) and chronic myeloid leukaemia cell line (Zhuang, Veltri and Long, 2019). One of these screens involved 33 cancer cell lines, many of which harbour active RAS mutations (Aguirre *et al.*, 2016). Another specifically looked at the protein interactomes of RAS isoforms with regards to proliferation behaviour changes (Adhikari and Counter, 2018). For RALB, 3 out of those 7 screens observed cell proliferation changes in a chronic myeloid leukaemia cell line (Liu *et al.*, 2019), glioblastoma cell line (MacLeod *et al.*, 2019) and a B-cell non-Hodgkin lymphoma cell line (Reddy *et al.*, 2017). Interestingly, RALB was also found to be a significant hit in a chronic myeloid leukaemia cell line that noted vesicle distribution had been affected (Lu *et al.*, 2018).

1.5 Aims and objectives

The RALGEF/RAL signalling pathway has emerged as being critically important in normal cell physiology. RAL proteins themselves have been implicated in many biological processes and whilst structurally and biochemically RALA and RALB are similar, they have been observed as having very distinct functions. Whether this is simply a reflection of spatially distinct interaction with the same set of effector proteins or that there are isoform specific effectors that remain to be discovered has yet to be determined. What's more, whilst RAL has been implicated in neoplastic cell growth and oncogenic RAS is known to be involved in the dysregulation of cellular growth and survival through MAPK signalling pathways, how and indeed if RAL also contributes to RAS-dependent tumourigenesis remains unclear. Furthermore, whilst a great deal has been learnt from studies on how

the RAL signalling network regulates biological and pathological processes, the approach of many has been to utilise knockdown techniques and/or overexpression of RAL, both of which have multiple drawbacks. Finally, although a few null and conditional RAL knockout mice have been generated, at the beginning of my project there were no available studies targeting and investigating RAL at the endogenous level in an appropriate human cell model. For this reason, isoform specific RAL knockouts, in a panel human colorectal cancer cell lines, harbouring different KRAS mutations will be generated. Tools needed to insert mutations at the endogenous loci, that express a constitutively active version of RAL, creating knock-in lines in the same cell panel will also be developed. These cell lines will be created using the CRISPR/Cas9 system and will allow the characterisation of RAL signalling at an endogenous level, thus build a detailed profile of RAL biology.

AIM:

To characterise isoform specific RAL function across a panel of colorectal cell lines.

OBJECTIVES:

- Quantify the absolute protein abundance contribution for each isoform.
- Build a detailed profile of the transcriptome and proteome involved in the RAL GTPase signalling network.
- Determine the phenotypic consequences of isoform specific RAL signalling.

Chapter 2: Materials & Methods

2.1 Cell Biology

2.1.1 Materials & Reagents

Dulbecco's Modified Eagle's Medium (DMEM) (#31966-021), Dulbecco's Modified Eagle's Medium/Nutrient Mixture F-12 Ham (DMEM/F-12) (#31331-028), McCoy's 5A Modified Medium (36600-021), Minimum Essential Medium (MEM) (#11095-080), Opti-MEM reduced serum medium (#31985-070), Trypsin 0.5% EDTA 10X (#15400-054) and Fetal Bovine Serum (FBS) (#10270106) were all purchased from Invitrogen (Paisley, UK). StemPro Accutase Dissociation Reagent (#A11105-01) was obtained from Fisher Scientific (Loughborough, UK). JetPEI transfection reagent (#101-01) was obtained from Polyplus Transfection (France). Lipofectamine LTX and PLUS reagent (#15338100) and Lipofectamine RNAiMAX transfection reagent (#10601435) were all purchased from Invitrogen. ON-TARGETplus Human RALA (5898) siRNA – SMARTpool, 5 nmol and ON-TARGETplus Human RALB (5899) siRNA – SMARTpool, 5 nmol were obtained from Dharmacon, Horizon Discover (Cambridge, UK). Puromycin dihydrochloride (#P7255) was purchased from Sigma-Aldrich, Merck (Poole, UK). All plasticware used for cell culture were obtained from Corning Inc. (NY, USA).

2.1.2 Cell culture

A panel of colorectal cancer cell lines were obtained for use in this study. Cell name, source and culture conditions are specified in Table 2.1. All cells were cultured in a 5% CO_2 atmosphere at 37°C. Confluent cells (~70%) were split every 3 days with an appropriate dissociation reagent, such as 1X trypsin-0.5% EDTA and ready to use StemPro Accutase (400-600 units/mL). Any subsequent gene-edited cell lines were cultured in the same way as the parental.

Table 2.1 Colorectal cancer cell line information.

Initial Cell Lines			
Cell name	**Source/ Catalogue No.**	**Media**	**Dissociation reagent/cell dilution**
CaCo2	ATCC/ HTB-37	MEM + 20% FBS	Accutase/ 1:3
DLD-1	ATCC/ CCL-221	RPMI + 10% FBS	Trypsin-EDTA/ 1:5
HCT116	In-house stock	DMEM +10% FBS	Trypsin-EDTA/ 1:10
LoVo	ATCC/ CCL-229	RPMI + 10% FBS	Trypsin-EDTA/ 1:5
SW48	In-house stock	McCoy's + 10% FBS	Accutase/ 1:3
SW403	ATCC/ CCL-230	DMEM/F-12 + 10% FBS	Accutase/ 1:5
SW620	ATCC/ CCL-227	DMEM + 10% FBS	Trypsin-EDTA/ 1:6
Generated Cell Lines			
HCT116 Parental clones	-	DMEM +10% FBS	Trypsin-EDTA/ 1:10
HCT116 RALA KO	-		
HCT116 RALB KO	-		
SW48 Parental clones	-	McCoy's + 10% FBS	Accutase/ 1:3
SW48 RALA KO	-		
SW48 RALB KO	-		

SW403 Parental clones	-	DMEM/F-12 + 10% FBS	Accutase/ 1:5
SW403 RALA KO	-		
SW403 RALB KO	-		
SW403 Cas9 inducible	-		
SW620 Parental clones	-	DMEM + 10% FBS	Accutase/ 1:5
SW620 RALA KO	-		
SW620 RALB KO	-		
SW620 Cas9 inducible	-		

2.1.3 DNA transfection

Different methods and reagents were tested to determine optimal transfection for both transient expression and gene editing in the chosen cell lines.

Lipid based reagents displayed similar transfection efficiency, with JetPEI and Lipofectamine LTX being the most effective. For transfections using JetPEI, cells were seeded in a 6-well plate at a density of $2x10^5$ and incubated overnight at 37°C. The following day cells were transfected using JetPEI following manufacturer's instructions. Each reaction contained 4 µL JetPEI in 96 µL 150 mM NaCl, added to a solution of 1 µg of DNA diluted in 100 µL NaCl and mixed gently by vortexing. Combined solutions were incubated for 20

minutes at room temperature and added dropwise to the cells. Plates were incubated at 37°C for 72 hours before harvesting.

For Lipofectamine LTX, cells were seeded in a 6-well plate at a density of $2x10^5$ and incubated overnight at 37°C. The following day cells were transfected using Lipofectamine LTX following manufacturer's instructions. Each reaction contained 6 µL LTX in 200 µl Opti-MEM, added to a solution of 1 µg of DNA and 2 µL PLUS reagent diluted in 200 µL Opti-MEM and mixed gently by vortexing. Combined solutions were incubated for 10 minutes at room temperature and added dropwise to 1800 µL of fresh medium on the cells. Plates were incubated at 37°C for 72 hours before harvesting.

2.1.4 siRNA knockdown

For a 40nM single siRNA knockdown experiment, cells were typically seeded at a density of $1.5x10^5$ cells per well in a 6-well plate and incubated overnight at 37°C. Cells were then transfected using Lipofectamine RNAiMAX according to manufacturer's instructions. Each reaction contained 2 µL siRNA (20 µL stock) (Table 2.2) in 83 µL Opti-MEM added to a solution of 2 µL RNAiMAX in 83 µL Opti-MEM. Combined solutions were incubated at room temperature for 20 minutes and added to 830 µL fresh medium in each well. Media was exchanged for 2 mL fresh medium after 6 hours at 37°C. For a double knockdown, RNAiMax was doubled to 4 µL and the Opti-MEM reduced so each solution is 85 µL in total.

Table 2.2 siRNA oligo sequences used for transfection.

siRNA		
	Target sequence	**Mol. Wt. (G/MOL)**
RALA	GGACUACGCUGCAAUUAGA	13,430.1
	CAAAUAAGCCCAAGGGUCA	13,430.1
	GAGGAAGUCCAGAUCGAUA	13,430.1
	GAAAUUCGAGCGAGAAAGA	13,415.1
RALB	GAAGAAGUUCAGAUAGAUA	13,385.1
	AAAGAAAGAUGUUGCUUAC	13,385.0
	GAAACAAGUCUGACCUAGA	13,415.1
	UCACAGAACAUGAAUCCUU	13,400.1

2.1.5 Antibiotic treatment

Transfected cells were treated with two separate concentrations of puromycin (1 and 2 µg/mL), 48 hours post-transfection to select the cells which had successfully taken up plasmids with a puromycin selection marker. Cell death was monitored over the following 4 days, with 70-80% cell death usually observed by day 2. If no visible cell death occurred after the second day, a second higher dose was added in fresh media. Any resultant colonies were grown and expanded into two new 6-well plates before single cell dilutions.

2.1.6 Virus production in HEK293T cells

For viral transduction, a combination of a retrovirus system and a 2nd generation lentiviral system was used depending on the transfer plasmid being utilised. The retroviral system involved a plasmid encoding the gene(s) of interest and the retrovirus packaging plasmid (pCL-10A1) that encodes both the packaging and envelope elements required for virus particle production. Whereas the 3 plasmids necessary for lentivirus production are one envelope plasmid encoding VSV-G (pMD2.G), one packaging plasmid (psPAX2 or psPAX2-D64V) and a lentiviral transfer plasmid encoding the insert of interest. (Table 2.3).

For any viral production, two 10 cm dishes were seeded with $3.5x10^6$ HEK293T cells and grown to reach 70% confluency. Culture medium was exchanged 4 hours prior to transient transfection (section 2.1.3).

For stable transfection and expression of WT and mutant RAL isoforms, a total of 2 µg of DNA (1 µg pCL-10A1 packaging vector, 1 µg pBabe-Puro transfer vector) was required per dish.

For the CRISPR/Cas9 knock-in cell lines, the psPAX2-D64V (Supplementary Supplementary Figure 7.1) is preferred over the psPAX2 plasmid. This plasmid generates integrase deficient lentiviral vectors due to a point mutation in the integrase gene. This stops any fluorescent markers from the transfer plasmid inserting into the target genome which would prevent any future assays involving fluorescent analysis. In total, 4 µg of transfer plasmid and 2 µg of each envelope and packaging plasmid plasmids was required per dish.

Following addition of transfection mixture, medium was changed the next morning and the supernatant harvested 48- and 72-hours post transfection. Viral particles were concentrated on day of harvest and were resuspended in PBS and stored at -80°C following flash freezing.

2.1.7 Viral Transduction

Six well plates were set up with $4x10^5$ cells. Polybrene at a final concentration of 8 µg/mL and 100 µL of unconcentrated virus were added to each well and the plates spinoculated for one hour at 1000 x g. Following incubation overnight at 37°C, medium was replaced, and the cells incubated for a further 48 hours (72 hours in total).

2.2 Molecular Biology

2.2.1 Reagents

Primers were ordered from Eurofins (Ebersberg, Germany) or IDT (Leuven, Belgium). OneTag Hot Start DNA Polymerase (#M0481S), Phusion High-Fidelity DNA Polymerase (#M0530S), Q5 High-Fidelity DNA Polymerase (#M0491S), Deoxynucleotide (dNTP) Solution Mix (#N0447S), Quick Ligation Kit (#M2200S), T4 Polynucleotide Kinase (#M0201S) and all restriction enzymes were purchased from New England Biolabs (NEB) (Hertfordshire, UK). Stable (#C3040I) and 5-alpha (#C2987I) competent cells were obtained from NEB. Auxotrophic E. *coli* AT713 bacteria were from The Yale Coli Genetic Stock Center (New Haven, USA). L-lysine-U-13C6-15N2 (#608041) and L-arginine-U-13C6-15N4 (#608033) and SOC medium (#S1797) was purchased from Sigma Aldrich-Merck (Poole, UK). Carbenicillin (#10163193), Ampicillin (#10419313) and TRIzol Reagent (#15596026) were purchased from Fisher Scientific (Loughborough, UK). QIAprep Spin Miniprep Kit (#27106), HiSpeed Plasmid Maxi Kit (#12663) and QIAquick Gel Extraction Kit (#28704) were from QIAGEN (Crawley, UK). Wizard™ SV Gel and PCR clean up system (#A9281) was purchased from Promega (Southampton, UK). Ultrapure Agarose powder (#15510-019) was obtained from Invitrogen (Paisley, UK) and the TAE buffer from National Diagnostics (Hull, UK). All DNA markers, 1kb (#N3232) and 100bp (#N3231) were purchased from NEB. All chemicals used, were obtained from Sigma Aldrich-Merck unless otherwise stated.

2.2.2 RNA extraction and integrity checks

Cells were grown to 70-80% confluency in 6 cm dishes. Cells were lysed using 1 mL Trizol directly into the dish before transferring cell lysate to 1.5 mL tube. In the fume hood, 200 µL of chloroform was added and each tube shaken vigorously for 15 seconds. The samples were incubated at room temperature for 3 minutes before centrifugation at 12,000 x g at 4°C for 15 minutes. The RNA present in the clear upper aqueous phase was pipetted into a new tube and incubated at room temperature for 10 minutes with 0.5 mL RNase free isopropanol. Samples were centrifuged at 12,000 x g at 4°C for 10 minutes. The RNA pellet was washed once with 1 mL 75% ethanol before centrifugation

at 7500 x g at 4°C for 5 minutes. The supernatant was removed by pouring and the RNA pellets left to air dry for 10 minutes before resuspension in 25-50 µL of RNase free H_2O. RNA concentration was quantified on a NanoDrop® 1000 Spectrophotometer and the integrity verified on an Agilent 2100 Bioanalyzer. Each sample was given an RNA integrity number (RIN) between 1-10, with 10 being of the highest quality. All samples scored 9-10/10 (data not shown) and were taken forward for RNA sequencing analysis.

2.2.3 RNA sequencing data and analysis

Following RNA extraction, all samples were sent to BGI Genomics (https://www.bgi.com/global/sequencing-services/rna-sequencing-solutions/transcriptome-sequencing/) to undergo RNA transcriptome sequencing using DNBSEQ technology. Initial analysis and bioinformatics were also carried out by BGI Genomics. Differentially Expressed Genes (DEGs) were detected with DEseq2 and Poisson distribution. DEseq2 was based on the negative binomial distribution, as described in (Love, Huber and Anders, 2014). Poisson distribution was performed as described in (Audic and Claverie, 1997).

Once this dataset was received from BGI, further evaluation of DEGs was carried out using a collection of data including the gene expression value, $\log_2$ fold change and adjusted p-value. This summary data file was filtered to show only those genes that were classed as a DEG (fold change ≥2.00 and adjusted P value ≤0.05). The $\log_2$ fold change was then used to generate a visual representation of gene expression changes for RAL KO vs parental samples for each cell line in the form of heatmaps. These were created using Cytoscape and Multiple Experiment Viewer (MeV) software.

2.2.4 Cloning

The following pBabePuro-RAL plasmids (Supplementary Supplementary Figure 7.2) were obtained for transient/stable isoform specific expression or RAL recombinant protein production. Plasmids pBabePuro-RALA(Q72L), pBabePuro-RALB(Q72L) and pBabePuro-RALB(G23V) were kindly sent from Professor Chris Counter at Duke University Medical Center, Durham. pBabe-

Puro-RALA-wt was a gift from William Hahn (Addgene plasmid #15251). pBabe-Puro-RALA-V23 was a gift from William Hahn (Addgene plasmid #15252). pBabe puro RALB wt was a gift from Channing Der (Addgene plasmid #19720).

2.2.5 Bacterial transformation

Ligations were transformed into competent E. *coli* via heat shock. Briefly, 5 µL of the ligation was added to 50 µL of bacteria and incubated on ice for 20 minutes. Bacteria were then heat-shocked for 1 minute at 42°C, then allowed to recover on ice for 2 minutes. Following this, 350 µL of SOC medium (Luria Broth with 10 mM $MgCl_2$, 10 mM $MgSO_4$, 2.5 mM KCl and 20 mM glucose) was added and incubated at 37°C, 245 rpm for 1 hour. Bacteria were then plated onto LB-Agar plates containing the appropriate antibiotic and grown at 37°C overnight.

Single colonies were grown in 5 mL LB broth supplemented with antibiotic overnight at 37°C, 245 rpm. Isolation of plasmid DNA was performed according to the manufacturer's instructions using the QIAprep Spin Miniprep Kit. Purified DNA was test digested and potential positive samples were sent for sequencing at the DNA Sequencing Service at University of Dundee, UK. Verified plasmids were expanded into 250 mL LB medium from glycerol stock scrapings and the DNA extracted using the Qiagen MaxiPrep Kit.

Table 2.3 Plasmid information.

Name	Product	Selection marker	Purpose
pBabe-Puro-RALA	WT RALA	Puromycin	Protein expression
pBabe-Puro-RALA-V23	G23V RALA	Puromycin	Protein expression
pBabePuro-RALA(Q72L)	Q72L RALA	Puromycin	Protein expression
pBabe puro RALB wt	WT RALB	Puromycin	Protein expression

pBabePuro-RALB(G23V)	G23V RALB	Puromycin	Protein expression
pBabePuro-RALB(Q72L)	Q72L RALB	Puromycin	Protein expression
ptrcHis-A-RALA	His-RALA (wt)	-	PSAQ
ptrcHis-A-RALB	His-RALB (wt)	-	PSAQ
pGEX4T3-GST-RalBD	GST-RALBD	-	RAL activity
LentiCRISPR(V2)-SFFV	gRNA, Cas9 enzyme	Puromycin	CRISPR/gene editing
pLeGO-EGFP/2A-puro	gRNA, HDR DNA template	Puromycin/ EGFP	CRISPR/gene editing
pCW-Cas9-Blast	Inducible Cas9 enzyme	Blasticidin	CRISPR/gene editing
psPAX2-D64V	Integrase deficient packaging plasmid	-	Lentivirus production
pMD2.G	VSVG envelope	-	Lentivirus production
pCL-10A1	Packaging vector	-	Retrovirus production
pEGFP-C1	EGFP	Neomycin	Transfection optimisation
*All plasmids are ampicillin resistant except LentiCRISPR(V2)-SFFV which has carbenicillin bacterial resistance and pEGFP-C1 which as kanamycin resistance.			

Table 2.4 A typical double restriction test digest reaction.

Component	Amount
10X Buffer	5 µL (1X)
Restriction Enzyme 1	1 µL
Restriction Enzyme 2	1 µL
DNA (1 µg)	X µL
DNase/RNase-free water	Up to 50 µL

2.2.6 Glycerol stocks

500 µL of 5 mL overnight bacterial culture was mixed with a 50% solution of glycerol and gently vortexed before being placed directly in -80°C for long-term storage.

2.2.7 Restriction digest

A typical test digest was set up as shown in Table 2.4. In general, 1 µg of DNA was used with the chosen enzyme/s (most restriction enzymes used were 20,000 units/mL) and the mixture incubated for 1 hour at 37°C.

2.2.8 Genomic DNA extraction

Following DNA transfection, cells were scraped, washed with PBS and pelleted by centrifugation at 1000 x g for 5 minutes. Cells for DNA isolation were lysed in 300 µL tail lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.1% sodium dodecyl sulphate) with 2 µL RNAase A (200 mg/mL) and 6 µL proteinase K (10 mg/mL) and incubated overnight at 55°C. Cell lysates were heated at 95°C for 10 minutes, followed by the addition of 1 µL glycogen (20 mg/mL) and 400 µL of isopropanol to precipitate the DNA. Samples were incubated on ice for 30 minutes before centrifugation for 20 minutes (14,000 x g). Resulting pellets were washed with 70% ethanol before finally being resuspended in nuclease free water.

2.2.9 Polymerase Chain Reaction (PCR) screen

DNA amplification by PCR was performed to identify correct gene editing by CRISPR/Cas9. OneTaq DNA polymerase was used to amplify the targeting

Exon 1 region in both RAL isoforms. Each PCR reaction was set up as shown in Table 2.5.

2.2.10 Agarose gel electrophoresis

Agarose gels were prepared with electrophoresis grade agarose and 1X TAE buffer (40 mM Tris-acetate, 1 mM Na_2-EDTA) and heated in a microwave until the agarose had fully dissolved, approximately 4 minutes. Ethidium Bromide was added at a final concentration of 0.5 µg/mL to allow visualisation of DNA using an ultraviolet light source. The mixture was then poured into a cassette and left to solidify at room temperature. Prior to loading, a 1:6 dilution of 6X gel loading dye, no SDS was added to each sample. 10 µL of either a 100 bp or 1 kb DNA ladder standard was loaded alongside the samples. DNA was resolved using 1X TAE buffer in an electrophoresis tank at 130 V for 30 minutes. DNA bands were confirmed under a uvidoc (uvitec).

Table 2.5 A typical PCR reaction for gene editing screen.

Component	Amount
OneTaq	5 µL
Forward and Reverse Primers (10 µM)	1 µL
DNA (200 ng)	1 µL
DNase/RNase-free water	8 µL

Table 2.6 Thermal cycler programme for general PCR.

	Step	No. of Cycles	Temperature	Time
1	Initial Denaturation	1	95°C	2 minutes
2	Denaturation	25	95°C	30 seconds
	Annealing		58°C	30 seconds
	Extension		68°C	30 seconds
3	Final Extension	1	68°C	10 minutes

2.2.11 Recombinant RALA & RALB standard production

To generate N-terminal His-tagged RALA and RALB recombinant proteins, WT RALA and WT RALB sequences were subcloned into the ptrcHis-A construct (Supplementary Supplementary Figure 7.3).

WT RALA and WT RALB were gel extracted from pBabe-Puro-RALA-wt (Addgene plasmid # 15251) and pBabe puro RALB wt (Addgene plasmid # 19720) following a two-hour restriction digest with *Bam*HI plus *Sal*I (RALA) and *Bam*HI plus *Xho*I (RALB) and subsequently ligated into a ptrcHis-A construct following digestion with the same enzymes.

Auxotrophic E. *coli* AT713 bacteria were used as this bacterial strain is suitable to generate isotope-labelled, full- length proteins due to its inability to synthesize arginine and lysine as a result of mutations present at *arg*A21 and l*ys*A21 (Taylor, 1970). Following ligation, all plasmids were transformed into AT713 bacteria and glycerol stocks (2.2.4) produced and stored at -80°C. Samples were sent for sequencing with a T7 forward (5′–TAATACGACTCACTATAGGG–3′) and T7 Terminator (5′–TATGCTAGTTATTGCTCAG–3′) primers.

2.3 CRISPR/Cas9 Gene Editing

For generation of knockout cells, a double gRNA approach was selected (Figure 2.1). This strategy uses a LentiCRISPR plasmid containing two gRNAs and the Cas9 enzyme.

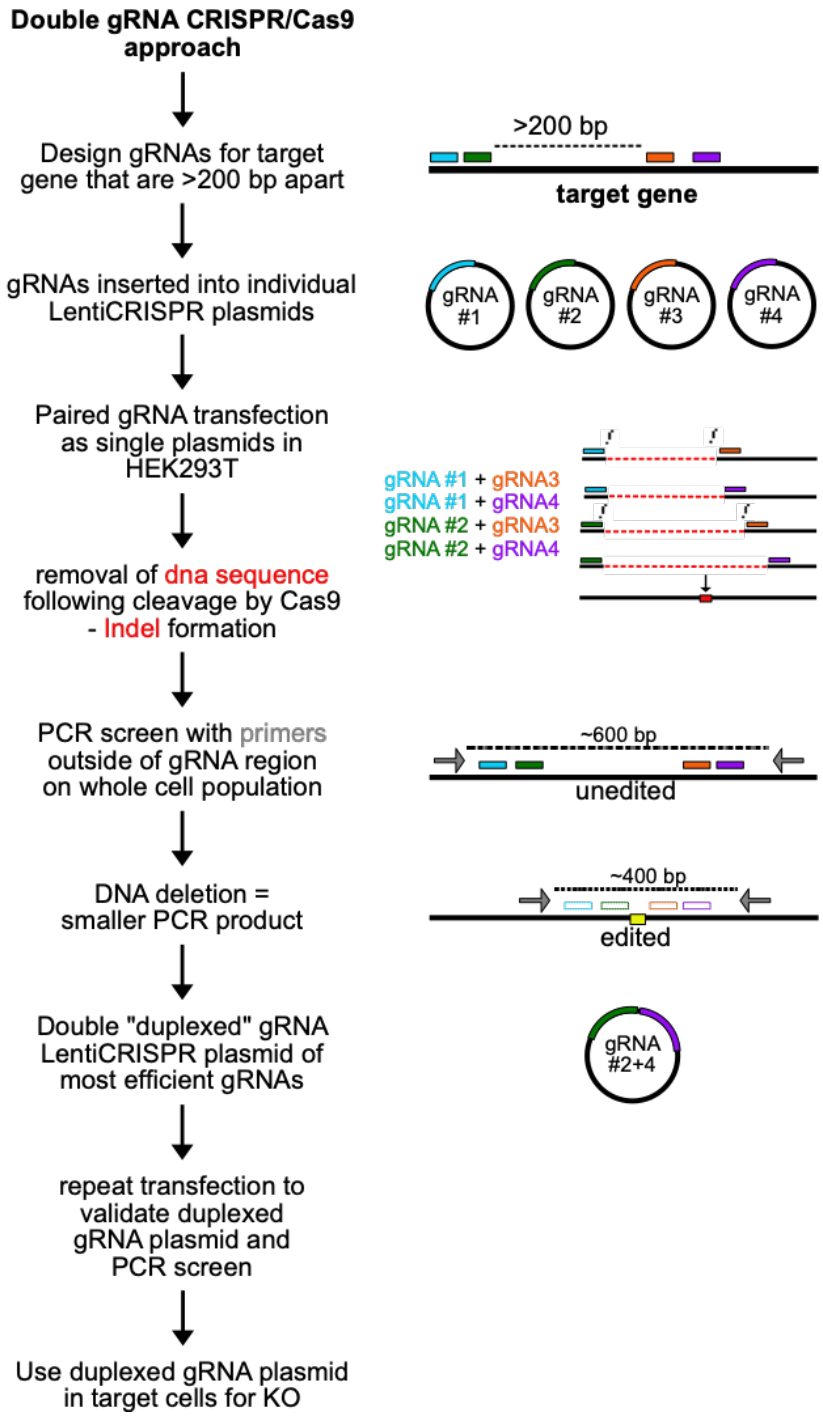


Figure 2.1 Strategy for generating isoform specific knockout cell lines.
A double guideRNA approach was used for generating knockout clones. This involved two gRNA inserted into one LentiCRISPR plasmid. Following targeted DNA cleavage by Cas9, the deletion of the sequence between the two gRNAs can be detected through the amplification of a smaller PCR product compared to the unedited cells.

2.3.1 Phosphorylation of guide oligonucleotides and ligation into LentiCRISPR

All gRNAs complimentary oligonucleotides with appropriate overhangs are shown in Table 2.7 with an additional G at the front of the 20-nucleotide sequence if absent.

Guide oligonucleotides were resuspended in DNase-free water at 100 µM and phosphorylated using T4 polynucleotide kinase at 37°C for 1h to allow for successful ligation. Oligonucleotides were annealed by heating at 95°C for 10 minutes, allowed to cool to room temperature, then diluted at 1:25 ready for ligation. For the ligation, 4 µL of the annealed oligonucleotides was mixed with 50 ng of *Bsm*BI digested LentiCRISPR(V2)-SFFV vector (Supplementary Supplementary Figure 7.4), 1 µL 10X T4 Ligase buffer, and 0.5 µL T4 Ligase, made up to 10 µL final volume. Ligations were incubated at 4°C overnight before transformation into Stable E. *coli* onto LB agar/carbenicillin plates (100 µg/mL).

Table 2.7 gRNA oligonucleotide sequences used for CRISPR/Cas9 gene editing.

gRNA No.	**RALA OLIGONUCLEOTIDES**	
	Forward (5'>3')	**Reverse (5'>3')**
1	CACC-GTAAAACAAGCCTATGACTCA	AAAC-TGAGTCATAGGCTTGTTTTA-C
2	CACC-GAATGAAGTATGCTCACTTAG	AAAC-CTAAGTGAGCATACTTCATT-C
3	CACC-GTTTACACAAAGTCATCATGG	AAAC-CCATGATGACTTTGTGTAAA-C
4	CACC-GAAAGTCATCATGGTGGGCAG	AAAC-CTGCCCACCATGATGACTTT-C
gRNA No.	**RALB OLIGONUCLEOTIDES**	
	Forward (5'>3')	**Reverse (5'>3')**
1	CACC-GTGTCATCAGCAGCTCTTCAG	AAAC-CTGAAGAGCTGCTGATGACA- C
2	CACC-GTCATCAGCAGCTCTTCAGT	AAAC-ACTGAAGAGCTGCTGATGA-C
3	CACC-GAAGGTGATCATGGTTGGCAG	AAAC-CTGCCAACCATGATCACCTT-C
4	CACC-GGTGATCATGGTTGGCAGCGG	AAAC-CCGCTGCCAACCATGATCA-C

2.3.2 Colony screening

Single colonies were picked and screened by polymerase chain reaction (PCR) using OneTaq DNA polymerase mix using a U6 promoter forward primer and the reverse primer for the corresponding guide. PCR positive colonies were identified by agarose gel electrophoresis and were cultured in 5 mL of LB supplemented with 100 µg/mL of Carbenicillin. Isolation of plasmid DNA was performed according to manufacturer's instruction using the QIAprep Spin Miniprep Kit. Clones were then sent for sequencing with a U6 promoter primer to the MRC PPU DNA Sequencing and Services in Dundee prior to larger scale DNA preparation (maxiprep).

Table 2.8 A typical PCR reaction for colony screen.

Component	Amount
OneTaq	5 µL
Forward and Reverse Primers (10 µM)	1 µL
DNase/RNase-free water	9 µL
Colony scraping	1

2.3.3 Assessing gRNA targeting and Cas9 cleavage.

Each gRNA targeting efficiency was assessed by co-transfection into HEK293T cells (section 2.1.3), the genomic DNA extracted (section 2.2.7) and screened for successful Cas9 cleavage, Indel formation and production of a smaller PCR product. Genotyping primer sequences are shown in Table 2.9 and one pair for each isoform selected following verification of efficiency (in this case that's Gen 1F + Gen1R for both isoforms). PCR amplification was set up and carried out as outlined in section 2.2.8.

Table 2.9 Genotyping primer sequences for PCR screen.

Genotyping Primers (sequence 5'-3')	
RALA	
Gen 1F (genotyping forward 1)	GCTGATACCCCATTATTGTCAGC
Gen 2F (genotyping forward 2)	GTCAGCATGCAGTTGTACTGGG
Gen 1R (genotyping reverse 1)	AATCAAGTTATGTACAGATGGCTC
Gen 2R (genotyping reverse 2)	CCATAAGTTTCCTGATTTGTAAAATC
RALB	
Gen 1F (genotyping forward 1)	GTGACTTCCAGAGTTGTTGG
Gen 2F (genotyping forward 2)	GAGGGCTCGCATGTCTGC
Gen 1R (genotyping reverse 1)	CGATGCAAGCAGAGATCTCTC
Gen 2R (genotyping reverse 2)	TAACAATGGAAACGATGCAAG

2.3.4 Duplexed gRNA LentiCRISPR cloning

The knockout cell lines generated require two gRNA for successful gene editing. Following validation of each gRNA by co-transfection as single LentiCRISPR plasmids, gRNA#2 and gRNA#4 for RALA and RALB were chosen for future experiments.

To ensure both gRNAs are taken up into the cell, a single LentiCRISPR plasmid containing both gRNAs is preferred over two separate single gRNA plasmids. To create this double or "duplexed" construct, gRNA#4 with adjoining U6 promoter from LentiCRISPR-gRNA#4 was amplified by PCR, with the addition of a linker sequence. This enabled ligation into the previously verified LentiCRISPR-gRNA#2 plasmid, resulting in a single plasmid with two gRNAs each with their own U6 promoter.

PCR amplification was carried out using Q5 High-Fidelity DNA Polymerase with forward primer 5′–CGGAATTCGTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTTTTAGCGCGTGCGCCAATTCTGCAGACAAATGGCT CTAGAGAGGGCCTATTTCCCATGATTCC–3′ and reverse primer 5′–GCGAATTCAAAAAAGCACCGACTCGG–3′. The PCR reaction was set up as show in Table 2.10 and run according to the thermal cycler programme described in Table 2.6. The amplified insert (U6 promoter + gRNA#4 + gRNA

linker) was cut from a 1.5% agarose gel using a Qiagen extraction kit, eluted and digested with restriction enzyme *Eco*RI for 2 hours at 37°C. A total of 4 µg of each RAL construct LentiCRISPR-gRNA#2 was digested with *Eco*RI for 2 hours at 37°C. The digested vector was gel extracted and purified using the Wizard™ SV Gel and PCR clean up system. The digested inserts were ligated into their corresponding RAL LentiCRISPR-gRNA#2 vectors overnight at 4°C before transformation into Stable competent cells. Positive colonies were identified as previously described and the isolated plasmid DNA transfected into HEK293T cells.

Table 2.10 A typical PCR reaction using Q5 high-fidelity DNA polymerase.

Component	Amount
Q5 polymerase	0.5 µL
dNTPs (10 µM)	1 µL
Forward and Reverse Primers (10 µM)	5 µL
5X buffer	10 µL
DNA (200 ng)	X µL
DNase/RNase-free water	X µL
total	50 µL

2.3.5 Doxycyclin inducible Cas9 cells

To generate inducible Cas9 cells, target cell lines were transduced (section 2.1.7) using a lentivirus containing the pCW-Cas9-Blast lentiviral vector (Supplementary Supplementary Figure 7.5). This construct contains the sequences for expression of doxycycline inducible Cas9 and constitutive expression of the selectable antibiotic marker, blasticidin.

Following positive selection with 1 µg/mL blasticidin for 5 days, cells were left to recover and expanded into duplicate 6-well plates before cell lysis to determine Cas9 expression by western blot. Cells were treated 24 hours with 8 µg/mL doxycycline plus 0.25% DMSO before protein extraction.

2.3.6 Insertion of gRNA#4 into pLeGO-EGFP/2A-Puro for knock-in

For generating knock-in cell lines, using a virus is the most effective way to ensure both CRISPR components (a single gRNA and HDR DNA template) get into a cell population (Figure 2.2). Since the Cas9 enzyme is already present in the chosen cell lines (section 2.3.4), gRNA#4 was subcloned from the verified LentiCRISPR(V2)-SFFV plasmids into a new pLeGO-EGFP/2A-Puro construct (Supplementary Supplementary Figure 7.6). This approach overcomes the problems associated with packaging the Cas9, gRNA and long HDR template into a single plasmid for virus production. It also offers the benefit of two positive selection markers, puromycin and enhanced GFP, to allow for easy detection of transduction efficacy.

PCR amplification of gRNA#4 for RALA and RALB was carried out using Q5 High-Fidelity DNA Polymerase with forward primer 5'-TAAGGTA CCGAGGGCCTATTTCCC-3' and reverse primer 5'-CTAGAGCTAGCGAAT TCAAAAAAAGCAC-3'. The PCR reaction was set up as show in Table 2.10 and run according to the thermal cycler programme described in Table 2.6. The amplified product was cut from a 1.5% agarose gel using a Qiagen extraction kit, eluted and digested with restriction enzyme *Kpn*I for 1 hour at 37°C. A total of 4 µg of the construct pLeGO-EGFP/2A-Puro (gifted from Dr Nicholas Harper, University of Liverpool) was digested with *Kpn*I and *Pme*I for 2 hours at 37°C. The digested vector was gel extracted and purified using the Wizard™ SV Gel and PCR clean up system. Each digested gRNA#4 insert was ligated into a pLeGO vector overnight at 4°C before transformation into Stable competent cells. Positive colonies were identified as previously described (2.3.4) and the isolated plasmid DNA transfected into HEK293T cells along with a LentiCRISPR-gRNA#2 to validate cutting efficiency. Transfection efficiency was also verified by GFP expression.

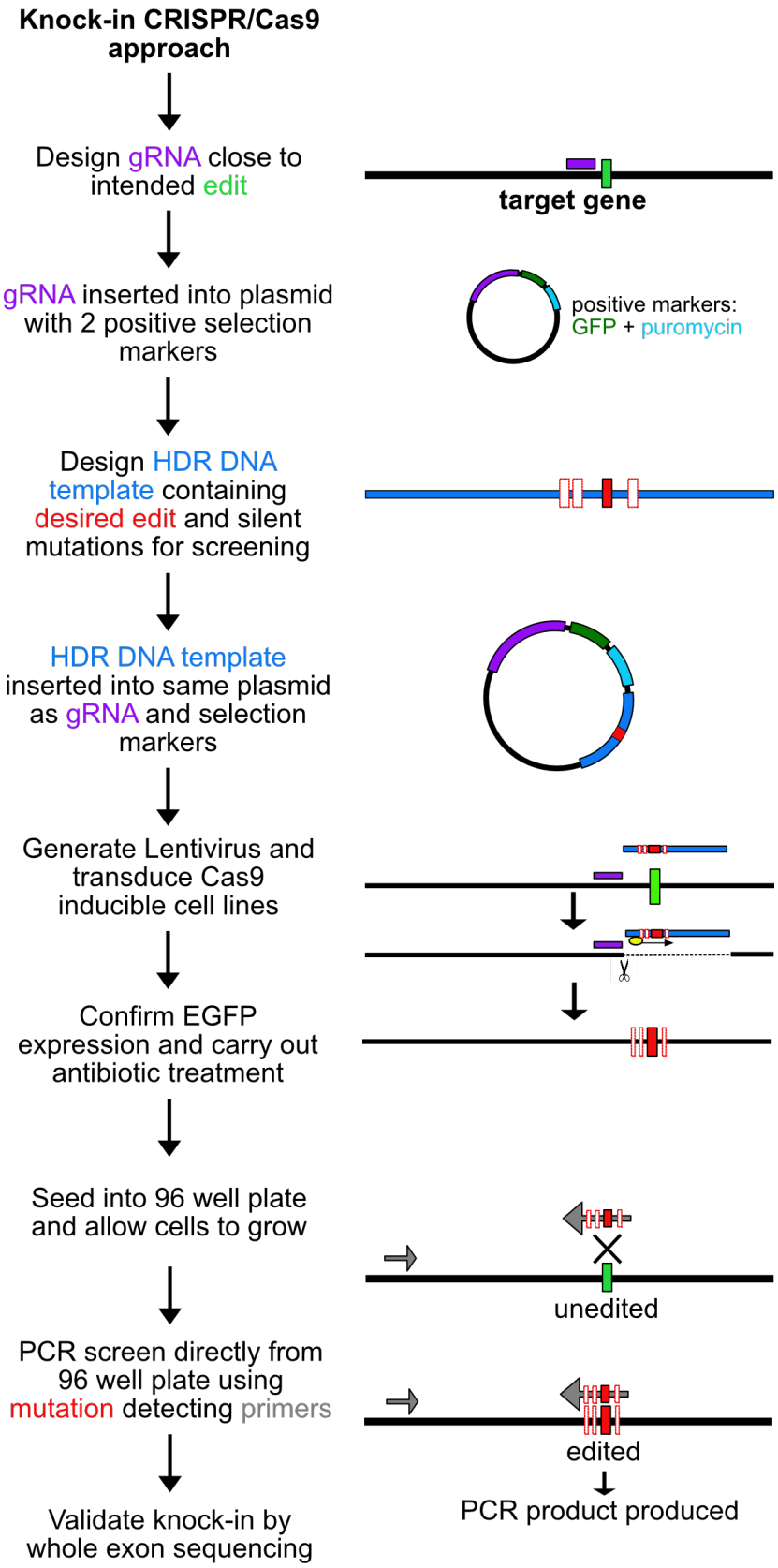
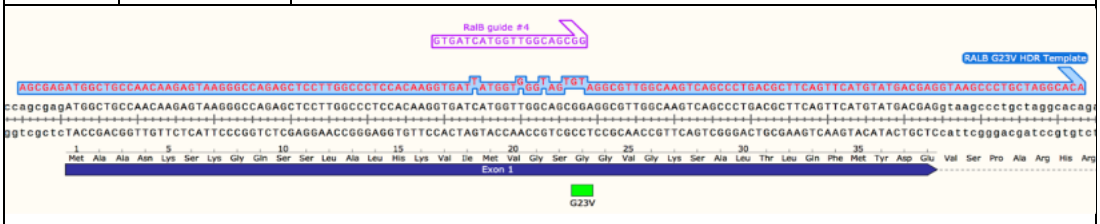


Figure 2.2 Strategy for generating knock-in cell lines.
Knock-in cell lines will be generated using a lentivirus approach. This will involve the transduction of target Cas9-inducible cells with a plasmid containing a single gRNA and an HDR template including the desired mutation and additional silent mutations. Following positive selection through GFP/puromycin, cells will be screened by PCR using mutant specific primers.

2.3.7 G23V HDR template

For homology directed repair to take place following a double stranded DNA break, as well as a gRNA, a DNA template including the desired mutation must be included (Table 2.11). For both RALA and RALB the glycine at amino acid 23 was changed for a valine to produce constitutively active RALA/B knock-in cell lines. As well as this specific mutation, other “silent” mutations were introduced into the template to aid the screening process following HDR.

Table 2.11 HDR DNA templates for RALA and RALB.

		Sequence 5'-3'
RALA	Template	…ATGGCTGCAAATAAGCCCAAGGGTCAGAATTCTTTGG CTTTACACAAAGTCATCATGGTGGGCTCCGTTGGAGTT GGCAAGTCAGCTCTGACTCTACAGTTCATGTACGATGA GGTAAGTGCTAATTTTATAATGGAT…
		
RALB	Template	…AGCGAGATGGCTGCCAACAAGAGTAAGGGCCAGAGCTCC TGGCCCTCCACAAGGTGATTATGGTGGGTAGTGTAGGCGT TGGCAAGTCAGCCCTGACGCTTCAGTTCATGTATGACGA GGTAAGCCCTGCTAGGCACA…
		

DNA templates were produced in 2 parts before ligation into the pLeGO-EGFP/2A-Puro that already contains gRNA#4 (section 2.3.5). Both parts were amplified using Q5 High-Fidelity DNA Polymerase with the reaction set up as outlined in Table 2.10, using primers specified in Table 2.12. Part 1 was amplified using normal PCR thermal cycler steps (Table 2.6), whilst part 2 was amplified by tail PCR (Table 2.13). Both parts were identified by agarose gel electrophoresis and extracted and eluted into 20 µL DNase/RNase-free water (section 2.2.9).

Using 15 µL of each template part, an overlap PCR reaction (Table 2.14) was set up with Q5 High-Fidelity DNA Polymerase and run for 15 cycles without primers (template parts will act as primers with each other). The PCR reaction was briefly paused and 5 µL of forward primer 5′–AAGGAAAAAAGCGGCCGCGTCAGCATGCAGTTGTACTGGG–3′ and reverse primer 5′–CTAGCTAGCCTTGTGATCTGCCTGTCTTGG–3′ for RALA and forward primer 5′–AAGGAAAAAAGCGGCCGCG AGGAGTAAAGCTAAAGAGCTGG–3′ and reverse primer 5′–CTAGCTAGCGCTTTACCTTGCCGCACG–3’ for RALB and run for another 20 cycles.

Table 2.12 Primers for amplification of HDR template parts 1 + 2.

	PART 1	
	Forward (5'>3')	**Reverse (5'>3')**
RALA	**HDR NotI** - AAGGAAAAAAGCGGCCGCGTCA GCATGCAGTTGTACTGGG	**HDR Overlap -** TGATGACTTTGTGTAAAGCCAAAG
RALB	**HDR NotI** – AAGGAAAAAAGCGGCCGCGAGG AGTAAAGCTAAAGAGCTGG	**HDR Overlap –** GCTCTGGCCCTTACTCTTG
	PART 2	
	Forward (5'>3')	**Reverse (5'>3')**
RALA	**HDR template –** ATGGCTGCAAATAAGCCCAAGG GTCAGAATTCTTTGGCTTTACA CAAAGTCATCATGGTGGGCTCC GTTGGAGTTGGCAAGTCAGCTC TGACCTACAGTTCATGTACGAT GAGGTAAGTGCTAATTTTATAA TGGAT	**HDR NheI -** CTAGCTAGCCTTGTGATCTGCCTG TCTTGG
RALB	**HDR template –** AGCGAGATGGCTGCCAACAAGA GTAAGGGCCAGAGCTCCTGGCC CTCCACAAGGTGATTATGGTGG GTAGTGTAGGCGTTGGCAAGTC AGCCTGACGCTTCAGTTCATGT ATGACAGGTAAGCCCTGCTAGG CACA	**HDR NheI -** CTAGCTAGCGCTTTACCTTGCCGC ACG
The HDR template for both isoforms were generated in two parts. Restriction sites and knock-in mutations are underlined.		

Table 2.13 Thermal cycler programme for tail PCR.

	Step	No. of Cycles	Temperature	Time
1	Initial Denaturation	1	94°C	3 minutes
2	Denaturation	13	94°C	20 seconds
3	Annealing		63°C	1 minute
4	Extension		72°C	1 minute
5	Denaturation		94°C	20 seconds
6	Annealing		61°C	1 minute
7	Extension		72°C	1 minute
8	Denaturation		94°C	20 seconds
9	Annealing		61°C	1 minute
10	Extension		72°C	1 minute
11	Denaturation		94°C	20 seconds
12	Annealing		50°C	1 minute
13	Extension		72°C	1 minute
14	Final Extension	1	72°C	5 minutes

Overlap PCR reactions were run on a 1% agarose gel and each complete template was extracted before digestion with *Nhe*I and *Not*I and subsequent ligation into the *Nhe*I/*Not*I pre-digested pLeGO-EGFP/2A-Puro.

To confirm successful ligation of the HDR DNA template into the vector, a PCR reaction was set up using 25 ng of plasmid DNA and 2 different primer pairs for double verification (Table 2.15). Correct plasmids were subsequently transformed into STABLE competent cells prior to larger scale DNA preparation and lentivirus production.

Table 2.14 Thermal cycler programme for overlap PCR.

	Step	No. of Cycles	Temperature	Time
1	Initial Denaturation	1	95°C	15 minutes
2	Denaturation	35	94°C	1 minute
	Annealing		55°C - 61°C	1 minute
	Extension		72°C	1 minute
3	Final Extension	1	72°C	10 minutes

Table 2.15 Primer Pairs for HDR DNA template ligation verification.

Primer Pair	**RALA**	
	Forward (5'>3')	**Reverse (5'>3')**
1	**HDR NotI** - AAGGAAAAAAGCGGCCGCGTCA GCATGCAGTTGTACTGGG	**G23V mutant -** GACTTGCCAACTCCAACGG
2	**G23V mutant -** GTGGGCTCCGTTGGAGTTG	**HDR NheI -** CTAGCTAGCCTTGTGATCT GCCTGTCTTGG
Primer Pair	**RALB**	
	Forward (5'>3')	**Reverse (5'>3')**
1	**HDR NotI** - AAGGAAAAAAGCGGCCGCGAGG AGTAAAGCTAAAGAGCTGG	**G23V mutant -** ACGCCTACACTACCCACC
2	**G23V mutant -** ATGGTGGGTAGTGTAGGC	**HDR NheI -** CTAGCTAGCGCTTTACCTTG CCGCACG
Knock-in mutations are underlined.		

2.3.8 PCR screen for knock-in

Due to time constraints, screening for knock-in clones was not carried out. However, the following protocol can be used to detect HDR template insertion and successful integration of mutations.

Following Cas9 induction by doxycycline treatment in conjunction with viral transduction (section 2.1.7) with either pLeGO-EGFP/2A-Puro-gRNA#4-RALA or pLeGO-EGFP/2A-Puro-gRNA#4-RALB, initial verification of successful transduction is confirmed through positive GFP expression. Transduced cells undergo antibiotic selection (section 2.1.5) before single cell isolation into a 96-well plate. After cell expansion and colonies formation, the plate is split into two new 96-well plates: one for knock-in cell line expansion and one for genotyping. Genomic DNA is extracted from one plate using a NaOH lysis method (Wang, Qi and Cutler, 1993). Briefly, 200 µL of 0.5 M NaOH is added to each well. Once mixed with the cells, 5 µL of the suspension is diluted in 495 µL 100 mM Tris-HCl, pH 8.0 and then 1 µL of the dilution is used as template DNA for PCR amplification.

To screen for a successful knock-in of the desired HDR template, isolated DNA undergoes genotyping PCR using primers situated in Exon 1 in combination with primers designed to bind to the mutated sequence only (Table 2.16). Once CRISPR/Cas9 gene edited cells have been identified and verified by PCR (and immunoblotting for knockout cell lines), they are sent off for sequence verification along with the genotyping primers listed in Table 2.17 that are situated in the introns at either end of exon 1.

Table 2.16 Genotyping primers for HDR amplification and Knock-in detection.

Primer Pair	**RALA**	
	Forward (5'>3')	**Reverse (5'>3')**
1	GCTGATACCCCATTATTGTCAGC	GACTTGCCAACTCCAACGG
2	GTGGGCTCCGTTGGAGTTG	AATCAAGTTATGTACAGATGGCTC
Primer Pair	**RALB**	
	Forward (5'>3')	**Reverse (5'>3')**
1	GTGACTTCCAGAGTTGTTGG	ACGCCTACACTACCCACC
2	ATGGTGGGTAGTGTAGGC	CGATGCAAGCAGCGATGTCTC
Knock-in mutations are underlined.		

Table 2.17 Genotyping primers for sequencing EXON 1.

Genotyping Primers (sequence 5'>3')	
RALA	
Gen 1F (genotyping forward 1)	AGATCCTTTAGCCTGCTTTCG
Gen 2F (genotyping forward 2)	CGTATCTGTGTGGACATATTGC
Gen 1R (genotyping reverse 1)	AATTTTAGTGACTTGGCCTGC
Gen 2R (genotyping reverse 2)	CATGCCAAACACACAATAAACG
RALB	
Gen 1F (genotyping forward 1)	GTGAGCATGTGTGTGAATGTGG
Gen 2F (genotyping forward 2)	CATGTGTGAGTGTGTTAGTGACC
Gen 1R (genotyping reverse 1)	GCCCACAGTAGGCTAAGTATCC
Gen 2R (genotyping reverse 2)	TAGCAAACAGCTCAAGCAGG

2.4 Protein Biochemistry

2.4.1 Reagents

Lysozyme (#L6876), PhosSTOP phosphatase inhibitor tablets (#04906837001) and all chemicals used were purchased from Sigma-Aldrich, Merck (Poole, UK) unless otherwise stated. Pierce™ BCA Protein Assay Kit (#23225), Pierce™ 660 nm Protein Assay Kit (#22662), full range rainbow molecular weight protein marker (#11580684) and GE Healthcare Glutathione Sepharose™ 4B Media (#11594935) were obtained from Fisher Scientific. NuPAGE 4-12% gradient Bis-Tris gels (10 well 1 mm #NP0321BOX, 15 well 1.5 mm #NP0303BOX), NuPAGE MES buffer (#NP0002-02), SimplyBlue SafeStain (#LC6060) were purchased from Invitrogen. Marvel skimmed milk powder was purchased from Premier brands, UK. Amersham™ nitrocellulose western blotting membrane (#GE10600002), GSTrap™ High Performance columns (#GE17-5281-01) and HisTrap™ High Performance columns (#GE17-5247-01) and Amicon® Ultra Ultra Centrifugal filters (#UFC900) were obtained from Sigma-Aldrich, Merck.

2.4.2 Cell lysis

Samples for isolation of protein were washed 3 times with ice cold 1X PBS and resuspended in 150 µL RIPA lysis buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% SDS, 1% sodium deoxycholate and 1% Triton X-100) containing 1X protease inhibitor cocktail and PhosSTOP phosphatase inhibitor cocktail. Plates were left rocking on ice for 10 minutes before lysates were collected into Eppendorfs and sonicated for 3 x 10 seconds. Samples underwent centrifugation at 13,000 g for 20 minutes at 4°C. Supernatants were transferred to a new tube and protein concentration quantified using a protein assay kit. Lysates were then stored at -20°C until use.

2.4.3 Concentration determination

The Pierce BCA protein assay kit was used to determine protein concentration of cell lysates. Bovine IgG was used for the standard curve concentrations. The kit was used according to manufacturer's instructions with a typical BCA protein assay depicted below. Samples were read on a Multiskan Spectrum plate reader at A_{562} following 30 minutes incubation at 37°C. The Pierce 660

nm protein assay kit was used to determine the concentration of recombinant proteins. A typical 660 nm protein assay is depicted below, and samples were read on a Multiskan Spectrum plate reader immediately after the addition of Reagent B.

Table 2.18 A typical BCA assay for protein concentration determination.

	IgG (µg/well)	**Water (µL)**	**RIPA Buffer (µL)**	**Cell Lysate (µL)**	**BCA Mixture* (µL)**
Standard Curve	- 2 4 6 8 10	13 11 9 7 5 3	2 2 2 2 2 2	- - - - - -	200
Sample	-	13	-	2	200
***BCA mixture was made with 50 parts BCA Reagent A with 1-part Reagent B.** **Final sample concentration (mg/mL) is calculated from the µg/well standard curve and divided by the volume of test sample.**					

Table 2.19 A typical 660 nm assay for protein concentration determination.

	Albumin Standard (mg/mL)	**Albumin Standard (µL)**	**Sample (µL)**	**Buffer (µL)**
Standard Curve	2 1.5 1 7.5 5 2.5 1.25 0	10 10 10 10 10 10 10 0	- - - - - - - -	- - - - - - - -
Sample	-	-	5	5
150 µL of Reagent B was added to each well				

2.4.4 Sodium Dodecyl Sulphate polyacrylamide electrophoresis (SDS-PAGE)

Protein samples were resolved by SDS-PAGE using precast NuPAGE 4-12% Bis-Tris gels. Protein lysates harvested in RIPA or NP-40 buffer were resuspended in 5X sample buffer (310 mM Tris-HCl pH 6.8, 50% w/v glycerol, 16% w/v 2-Mercaptoethanol, 15% w/v SDS, 0.05% w/v Bromophenol Blue) and boiled at 98°C for 5 minutes. Bacterial pellets or "hot lysates" were resuspended in 10X hot lysis buffer (1M DTT, 1% Bromophenol Blue) and boiled at 110°C for 5 minutes before loading onto the gel. Depending on the concentration of cell lysate collected, 10-20 µg of protein was usually loaded into each well.

Precast NuPAGE 4-12% gradient Bis-Tris gels in a XCell SureLock Mini-Cell System (10, 15 and 20 wells) were used for most experiments. MES buffer was used to separate lower molecular weight proteins with the gels run at 150 V for 90 minutes.

2.4.5 Western blot

Following electrophoresis, proteins were transferred to a nitrocellulose membrane for 1 hour at 100 V. Each membrane was blocked with 5% milk in 1X TBST (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween 20) for 2 hours at room temperature. Blots were washed three times with 1X TBST and incubated with primary antibodies (Table 2.20) diluted in 5% milk/ 1X TBST overnight at 4°C. Unbound antibodies were washed off 3 times with 1X TBST, changing the TBST every 5 minutes. Membranes were then incubated with a secondary antibody (Table 2.21) for 45 minutes at room temperature. Membranes were washed with 1X TBST for 15 minutes as previously described and results visualised on an Odyssey CLx LI-COR.

Quantification of protein expression was carried out using Image Studio Lite software and statistical analysis using Welch's Analysis of Variance (ANOVA) parametric test was conducted on GraphPad PRISM software. The Welch's ANOVA was selected to account for the lack of homogeneity of variance within sample groups.

Table 2.20 Primary antibodies for western blotting.

Antibody Name	Target Protein	Species	Source	Dilution
Beta Actin	Actin	M	Protein Tech, 60008-1	1:10,000
Phospho-Aurora A (Thr288)/Aurora B (Thr232)/Aurora C (Thr198) (D13A11)	Phospho-Aurora A/B/C	R	CST, 2914	1:1000
Cas9 (7A9-3A3)	Cas9	M	CST, 14697	1:1000
DUSP4 [EPR19881]	DUSP4	R	Abcam, ab216576	1:1000
DUSP5 [EPR19684]	DUSP5	R	Abcam, ab200708	1:1000
EphA2 (D4A2)	EphA2	R	CST, 6997	1:1000
Phospho-EphA2 (Ser897) (D9A1)	Phospho-EphA2 (Ser897)	R	CST, 6347	1:1000
Phospho-EphA2 (Tyr772)	Phospho-EphA2 (Tyr772)	R	CST, 14069	1:1000
Exoc2 [EPR9420]	Sec5	R	Abcam, ab140620	1:1000
Exoc8	Exo84	R	Abcam, ab106121	1:500
KRAS	KRAS	M	LSBio, C175665	1:2000
His Tag	His	M	Amersham, 27471001	1:3000
P44/42 MAPK (Erk1/2) (137F5)	ERK	R	CST, 4695	1:1000
Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) (D13.14.4E)	Phospho-ERK	R	CST, 4370	1:2000
Phospho-PKCa/b II (Thr638/641)	Phospho-PKCa/b II	R	CST, 9375	1:1000
RALA	RALA	R	CST, 3526	1:1000
RALB	RALB	R	CST, 3523	1:1000
RALBP1	RALBP1	R	Abcam, ab33446	1:500
RALGDS	RALGDS	R	Abcam, ab65204	1:1000
WEE1	WEE1	R	CST, 4936	1:1000
(M: mouse; R: rabbit; CST: Cell Signalling Technology).				

Table 2.21 Secondary antibodies for western blotting.

Secondary Antibody	Catalogue No.	Dilution
IRDye® 680RD Donkey anti-Rabbit	926-68073	1:15,000
IRDye® 680CW Donkey anti-Mouse	926-68072	1:15,000
IRDye® 800CW Donkey anti-Rabbit	926-32213	1:15,000
IRDye® 800CW Donkey anti-Mouse	926-32212	1:15,000

2.4.6 GST-RALBD protein production

Construct pGEX-GST-RALBD was a gift from Dr Ignacio Rubio. GST-RALBD contains amino acids 397-518 of the binding domain of RAL effector RALBP1 (RAL binding protein/RLIP76). Following transformation into DH5-alpha competent cells, single colonies were grown overnight in 5 mL LB media with 100 µg/mL Ampicillin and glycerol stocks made the following day.

Scrapings from each glycerol stock were grown separately in 5 mL LB broth for 6 hours at 37°C at 245 rpm. 100 µL of each culture was transferred to 100 mL of LB broth plus ampicillin and grown overnight under the same conditions. All the overnight culture was spiked into 900 mL of fresh LB broth and grown to an OD600 of 0.6, before induction with isopropyl B-D-1thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Bacteria were incubated for overnight at 20°C, and the bacteria collected by centrifugation at 3000 x g, 15 minutes. Pellets were washed once with ice cold 1X PBS and pelleted again before storing at -80°C.

2.4.7 GST-RALBD protein purification

Bacterial pellets were resuspended in ice- cold IMAC20 (20 mM Tris-HCl, pH 7.4, 0.5 M NaCl) 5 mL per 1 g of bacteria, supplemented with bacterial protease inhibitors (1:500, #P8465, Sigma-Aldrich). Lysozyme was added to a final concentration of 1 mg/mL and incubated on ice for 30 minutes with regular agitation. The bacterial lysate underwent sonication 5 x 10 seconds with 30 second intervals on ice, ahead of centrifugation at 82,000 x g for 30 minutes at 4°C. The supernatant was collected and filtered using a 0.2 µm syringe filter before purification on the ÄKTA purifier system equipped with Frac950 and UPC900 (GE Healthcare, Buckinghamshire, UK) using a GSTrap HP columns according to manufacturer's instructions. Data was analysed on Unicorn 5.31 software.

Table 2.22 Buffers for GST-protein purification.

GST-RALBD purification buffers for ÄKTA		
	Reagent	**Final Concentration**
Buffer A (Binding Buffer)	NaH_2PO_4	10 mM
pH 7.4	NaCl	0.14 M
Buffer B (Elution Buffer)	Tris-HCl	50 mM
pH 8.0	Reduced glutathione	10 mM

2.4.8 Protein concentration and storage

Purity of GST-RALBD protein was checked by subjecting 20 µL from each elution well to SDS-PAGE, using 4-12% bis-tris NuPAGE gel, before colloidal blue staining.

GST-RALBD protein was concentrated and eluted using an Amicon Ultra 30-kDa MWCO filter at 3000 x g with wash and store buffer (10% glycerol, 50 mM Tris (pH 7.4), 150 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT) until the total volume reduced to ~0.5 mL. Concentration was determined by BCA assay and Pierce 660 nm (Table 2.18 and Table 2.19) before being flash frozen in liquid nitrogen and stored at -80°C in LoBind tubes.

2.4.9 RAL activity assays

Cell Preparation and Lysis

PAR and RAL KO cell lines were seeded in 10 cm dishes to be at least 80% confluent for all activity assays. Following overnight serum starvation, all media was exchanged for fresh medium supplemented with 10% FBS and 50 ng/mL EGF to stimulate the RAL pathway. Treatment time was determined following an EGF time course (0-20 minutes) on a parental clone from each cell type. Dishes were immediately placed on ice, washed once with ice cold 1X PBS and 150-300 µl of cell lysis buffer (25 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 0.25 M NaCl and 2% Igepal CA-630 (Nonidet P-40) was added to each dish for at least 30 seconds before scraping into a 1.5 mL eppendorf. Cell lysates underwent centrifugation at 1000 x g for 2 minutes at 4°C before the supernatant was transferred to a new LoBind Eppendorf and flash frozen in liquid nitrogen. A small volume of lysate was left with the pellet for use in a BCA assay to determine concentration. Treated cell lysates were stored at -80°C until use.

RALA/B Pulldown

Previously treated lysates were defrosted in a water bath at room temperature and made up to 1 mg/mL in cell lysis buffer. 400 µl (400 µg) of each lysate was added to a new 1.5 mL tube already containing 10 µg of Glutathione Sepharose beads pre-mixed with 10 µg of GST-RALBD in 200 µL cell lysis buffer (600 µL total in tube). Samples were incubated with rotation at 4°C for 1 hour. Following centrifugation at 1000 x g for 2 minutes at 4°C, supernatant was carefully removed with gel loading tips to avoid disturbing the beads and washed once with wash buffer (25 mM Tris-HCl pH 7.5, 30 mM $MgCl_2$, 40 mM NaCl) and resuspended in 16 µL 4X loading buffer and heated at 95°C for 5 minutes. 8 µL was loaded in duplicate onto a 4-12% Bis-Tris gel for separate RALA and RALB probing by western blot. 20 µg of total cell lysate was run alongside each pulldown to probe for total RALA/B and actin. Quantification of active RALA and RALB was performed using Image Studio Lite software and normalised to actin expression across 3 separate repeats.

2.4.10 Reverse Phase Protein Array (RPPA) analysis

Cell pellet samples were collected and sent to the Functional Proteomics RPPA Core Facility at the MD Anderson Cancer Center (https://www.mdanderson.org/research/research-resources/core-facilities/functional-proteomics-rppa-core/rppa-process.html). Cell pellets were lysed using RPPA buffer and protein extracted. Lysates were serially diluted in 5 two-fold dilutions and using an Aushon Biosystems 2470 arrayer, applied onto nitrocellulose-coated slides. The slides were probed with 466 previously validated primary antibodies followed by detection with secondary antibodies. Samples were then visualised by DAB colourimetric reaction to produce stained slides and scanned on a Huron TissueScope scanner. Primary data analysis conducted at the facility generated normalised linear data values for protein expression and modification. Once this dataset was received, a correlation analysis was conducted between cell lines. A correlation coefficient value was generated using the CORRELL function in Microsoft excel. Perfect correlation between two samples is indicated by a coefficient of +1, whereas -1 indicates an inverse relation and lack of expression similarity.

To determine the expression fold change for KO vs PAR, the normalised linear values for each RAL status within a cell line were averaged (PAR1+PAR2, AKO1+AKO2, BKO1+BKO2) and a ratio for AKO/PAR and BKO/PAR were generated. Each ratio was subsequently turned into a log ratio for visualisation of fold change.

To allow for comparison of isoform specific knockout verses parental protein expression, $\log_2$ fold change of ±0.5 were used to generate heatmaps in Cytoscape and Multiple Experiment Viewer (MeV) software.

2.5 Protein Standard Absolute Quantification (PSAQ)

The PSAQ method described below was optimised throughout my project and adapted from: Hood FE, Sahraoui YM, Jenkins, RE and Prior IA. Absolute quantitation of GTPase protein abundance. *Methods in Molecular Biology* (2021) in press.

2.5.1 Production of heavy labelled His tagged RAL standards

Scrapings from ptrcHis-A-RALA and ptrcHis-A-RALB glycerol stocks were grown separately in 5 mL LB broth for 6 hours at 37°C at 245 rpm. 100 µl of each culture was transferred to 100 mL of 1X M9 minimal media (Table 2.23) and grown overnight under the same conditions. All the overnight culture was spiked into 900 mL of fresh 1X M9 minimal media and grown to an OD600 of 0.6, before induction with isopropyl B-D-1thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Bacteria were incubated for a further 3 hours, and the bacteria collected by centrifugation at 3000 x g, 15 minutes. Pellets were washed once with ice cold 1X PBS and pelleted again before storing at -80°C.

2.5.2Affinity purification using His-Trap columns

Bacterial pellets were prepared as outlined in section 2.4.7 before purification on the ÄKTA purifier system using the following steps:

1. Wash the His-Trap column at 1 mL/min with 5 column volumes (CVs) of Binding Buffer (20 mM Tris-HCl, pH 7.4, 0.5 M NaCl, 20 mM Imidazole), then 5 CVs of Elution Buffer (20 mM Tris-HCl, pH 7.4, 0.5 M NaCl, 500 mM Imidazole)
2. Equilibrate the His-Trap column with 10 CVs of Binding Buffer.
3. Batch load the filtered lysate supernatant onto the HisTrap column through a 5 mL loop using a 5 mL syringe.
4. Equilibrate the His-Trap column with 10-15 CVs of Binding Buffer.
5. Run the imidazole gradient:
    a. Monitor protein elution by reading absorbance at 280 nm using a UV module throughout
    b. 1 mL/min flow rate throughout
    c. 0-60% Elution Buffer (20-300 mM imidazole) over 21.5 CVs
    d. 0.25 mL elution fractions

e. 3 mLs gradient delay

2.5.3 Purification check

For the filtered lysate and flow-through samples (40 µL each), 40 µL of hot lysis buffer (50 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol) was added and the samples boiled at 110°C for 15 minutes with frequent vortexing. Following microfugation for 5 minutes at 17,000 x g, 10 µL of 10X hot lysis buffer gel loading dye (1 M DTT, 1% Bromophenol blue) was added before a second boil at 95°C for 5 minutes.

20 µL from each elution fraction within an observed UV peak was added to 5 µL of 5X SDS-PAGE sample buffer before boiling at 95°C for 5 minutes.

To check for successful enrichment and purity of the His-tagged RALA/proteins, 10 µL of each was run per lane on replicate 4-12% Bis-Tris NuPAGE gels using 1X MES running buffer. One gel was stained using SimplyBlue SafeStain and the other was transferred onto membrane for probing with an anti-His tag and RALA and RALB antibodies (see 2.4.5).

2.5.4 Size-based purification using gel filtration

Appropriate fractions were collected and pooled together before concentration in gel filtration buffer (40 mM Tris-HCl, pH 7.4, 50 mM NaCl, 5 mM $MgCl_2$) using an Amicon Ultra 15 kDa MWCO filter column, under centrifugation at 3000 x g until the total volume reached ~500 µL.

Concentrated proteins were run through a Superdex 200 column previously equilibrated with 1.5 CVs of gel filtration buffer and collected in 1.25 CVs with an elution fraction size of 0.25 mL at a flow rate of 0.5 mL/min throughout.

Again, 20 µL of each fraction of interest (based on the peak observed in UV trace) was taken and prepared as before to confirm purity using SDS-PAGE and coomassie staining/western blotting.

Table 2.23 1X M9 minimal media recipe.

1X M9 minimal media pH 7.4 (1L)		
	Components	**Final Concentration**
5X M9 media pH 7.4 (autoclaved)	120 mM Na_2HPO_4 55 mM KH_2PO_4 21.5 mM NaCl 10 mM NH_4Cl	1X
Hydrophobic amino acids (2 mg/mL each)	L-Alanine, L-Isoleucine, L-Leucine, L-Methionine, L-Phenylalanine, L-Proline, L-Tryptophan, L-Valine	200 µg/mL
Hydrophilic amino acids (0.5 mg/mL each)	L-Aspartic Acid, L-Asparagine, L-Cysteine, L-Glutamic acid, L-Glutamine, Glycine, L-Histidine, L-Serine, L-Threonine, L-Tyrosine	100 µg/mL
Heavy (light for tests) L-Arginine and L-Lysine (100 mg/mL each)	L-arginine-U-$^{13}C_6$-$^{15}N_4$ L-lysine-U-$^{13}C_6$-$^{15}N_2$	200 µg/mL
glucose	glucose	0.4%
Thiamine-HCl (1 mg/mL)	Thiamine HCl solution	1 µg/mL
$MgSO_4$ (1 M)	$MgSO_4$ solution	1 mM

$CaCl_2$ (1M)	$CaCl_2$ solution	0.1 mM
$ZnSO_4$ (0.1 M)	$ZnSO_4$ solution	20 µM
Trace metals mix	Trace Metals A5 with Co	0.1%
Ampicillin	Ampicillin solution	100 µg/mL

2.5.5 Storage and concentration determination of standards

Desired gel filtration fractions were combined and concentrated using 10 kDa molecular weight cut-off 0.5 mL Centricon columns until the final volume of the protein standard is 150-400 µL. RAL isoform concentration was checked by BCA assay and the protein diluted using using 5X SDS-PAGE sample buffer to a final concentration of 1X sample buffer. Samples were boiled for 5 minutes at 95°C before being aliquoted into LoBind Eppendorfs and flash frozen in liquid nitrogen and stored at -80°C.

The percentage of heavy labelling was determined using standard LC-MS or MRM by comparing heavy protein to a known concentration of light His-RALA and analysed using ProteinPilot5 software.

2.5.6 Harvesting cells

Cells were grown in 10 cm dishes to 80% confluence before dissociation with the appropriate reagent outlined in section 2.1.2. Cells were pelleted by centrifuging at 150 x g for 5 minutes and the supernatant aspirated off before resuspension of the cell pellet in 9 mL of medium supplemented with 10% FBS. A total of 9 mL was immediately transferred to a fresh 15 mL centrifuge tube for collecting the cell pellet (section 2.5.7). The remainder of the cell suspension was used for counting on a haemocytometer, with an average of two counts being within 5% of each other.

2.5.7 Collecting cell pellets

The 9 mL of cell suspension was centrifuged at 150 x g for 5 minutes at room temperature. The media was aspirated off and the cell pellet resuspended in 10 mL of ice-cold PBS. Following a repeat centrifugation at 4°C, each cell pellet was resuspended in 300 µL of ice-cold PBS using a wide bore tip and transferred to a LoBind Eppendorf on ice. A fresh tip was used to add another

300 µL ice-cold PBS to the tube and the sample gently mixed before spinning at 150 x g for 5 minutes at 4°C. Each pellet was resuspended in 300 µL of ice-cold PBS and transferred to a new LoBind Eppendorf before centrifugation. All PBS was removed using gel loading tips and either flash frozen in liquid nitrogen and store at -80°C or lysed on ice in NP40 lysis buffer.

2.5.8 Cell lysis and spike-in

Cell pellets were resuspended in 140-300 µL of NP40 lysis buffer and incubated on ice for 10 minutes. Following centrifugation at 17,000 x g for 15 minutes at 4˚C, supernatant was collected and transferred to a new LoBind Eppendorf.

Protein concentration was determined by BCA assay and final samples of 20 µg of cell lysate with 2 ng of heavy His-RALA and His-RALB each were made and prepared for SDS-PAGE.

Samples were separated by SDS-PAGE on a self-poured 15 well, 1 mm 10% gel at 100 V until all rainbow protein markers could just be visualised. Alternating lanes of rainbow marker and samples were run to aid in ease of gel extraction and reduce sample contamination.

2.5.9 In-gel Digestion

Once samples have been separated by SDS-PAGE, the gel was fixed with 10% Acetic acid / 50% methanol for 10 minutes at room temperature with gentle rocking. Gels were washed in HPLC grade water 3 x 5 minutes, moved to a laminar flow hood and the region of interest (17-31 kDa) excised using sterile scalpel blades which were cut further into approximately 1 mm cubes and placed into a LoBind Eppendorf.

Samples then underwent the following in-gel digestion steps:

1. Dehydrate samples by adding 300 µL of 100% ACN to cover the gel pieces. Incubate at 37°C for 10 minutes, 900 rpm on a thermoshaker.
2. Dry tubes completely using a SpeedVac for 5 minutes at 37°C.
3. Reduce samples by adding 300 µL 10 mM DTT and incubate at 56 °C for 1 hour at 900 rpm on a thermoshaker. Discard supernatant.

4. Alkylate samples by adding 300 µL of 50 mM IAM. Incubate in the dark for 30 minutes, at room temperature, 900 rpm on a thermoshaker. Discard the supernatant.
5. Wash samples by adding 300 µL of 100 mM Ambic. Incubate for 15 minutes at room temperature, 900 rpm on a thermoshaker. Discard the supernatant.
6. Wash samples by adding 300 µL of 50 mM Ambic / 50% ACN. Incubate for 15 minutes at room temperature, 900 rpm on a thermoshaker. Discard the supernatant.
7. Dehydrate the gel slices by adding 300 µL of 100% ACN. Incubate for 5 minutes at room temperature, 900 rpm on a thermoshaker. Discard the supernatant. Repeat this step as many times as necessary until the pieces are small, hard and opaque white.
8. Dry tube completely by SpeedVac for 5-10 minutes at 37°C.
9. Make up a single stock of 2.5 ng/µL trypsin to cover all samples (typically 60-90 µL per sample), plus excess, in reaction buffer.
10. Add 1 volume trypsin to each tube and incubate for ~16 hours at 37°C.
11. Add 1 volume of 100% ACN. Incubate for 30 minutes at 30°C, 900 rpm on a thermoshaker.
12. Transfer the supernatant to a fresh LoBind Eppendorf. Repeat step 15 if pieces are not completely dehydrated.
13. Add 1 volume of freshly prepared 1% formic acid to gel pieces. Incubate at room temperature for 20 minutes at 900 rpm on a thermoshaker. Transfer the supernatant to the same Eppendorf as in step 16.
14. Repeat step 17.
15. Add 1 volume of 100% ACN to the gel pieces. Incubate at room temperature for 10 minutes at 900 rpm on a thermoshaker. Transfer the supernatant to the same Eppendorf as in step 16. Repeat until gel pieces shrink and turn white.
16. SpeedVac the peptides overnight at 37°C to dry them and store at -20°C.

2.5.10 Desalting and clean-up

Once peptides were dry, each sample was resuspended in 100 µL 0.1% TFA before microfugation at 12,000 x g for 5 minutes. Samples were transferred to

prelabelled glass vials and run on the Agilent 1260 to desalt. The peptides were then dried overnight at 37°C in the SpeedVac.

2.5.11 Quantification of endogenous RAL levels in cell lysates

Desalted samples were reconstituted in 11 µL of 0.1% formic acid and delivered into the mass spectrometer in 5 µL aliquots.

The precise molecular weight (M_w) (in Daltons, Da) of the heavy standard (including His-tag) and was obtained by entering the protein sequence into ProtParam on the ExPasy webpage. The number of moles was then calculated by dividing the mass (g) by the molecular weight. A ratio of endogenous(light)/standard(heavy) AUC was calculated and multiplied by the number of moles to give the number of moles per lane. The calibration curve equation for each peptide can also be used to determine the number of moles of protein per lane. The following quantifications were then calculated:

1. The number of moles per µg lysate (moles per lane divided by µg lysate per lane).
2. The number of cells per µg lysate (number of cells in the pellet divided by the total µg lysate yielded).
3. The number of moles per cell (step 1 divided by step 2).
4. The number of molecules of RALA and RALB per cell by multiplying step 3 by Avogadro's number.

Chapter 3: Selecting cells for RAL biology analysis

3.1 Introduction

RAL is known to play a part in many cellular functions, including the exocyst and exocytosis (Moskalenko *et al.*, 2002), actin organisation (Sugihara *et al.*, 2002) and gene transcription (Frankel *et al.*, 2005). Studies like these have predominantly relied on the use protein overexpression to highlight the biological roles of RAL, with the caveat that overexpression perturbs cell signalling pathways. Furthermore, outside of its immediate effectors, the RAL pathway and its functions remains poorly understood in comparison to other RAS effector pathways.

Endogenous RAL has been looked at using siRNA/shRNA and has provided evidence to support RAL's participation in mitochondrial fission (Kashatus *et al.*, 2011), invadopodia formation (Neel *et al.*, 2012), autophagosome assembly (Bodemann *et al.*, 2011), and the transformation and growth promoting activity of RAS (Chien and White, 2003; Kidd *et al.*, 2010; Guin *et al.*, 2013). However, knockdown experiments have their own limitations that include silencing of sequence specific non-target mRNAs and impairment of endogenous microRNA function (Boettcher and McManus, 2015).

Protein overexpression has supported the idea that RAL plays a role in cancer. It has been demonstrated that the presence of a negative RALA mutant blocks overproduction of urokinase-type plasminogen activator and metalloproteases leading to v-Src and v-RAS induced tumourigenicity and an invasive, metastatic phenotype (Aguirre-Ghiso *et al.*, 1999). Activation of c-Src by EGF has been shown to occur through a RAL dependent mechanism with constitutively active RAL leading to phosphorylation of Stat3 and cortactin (Goi *et al.*, 2000).

However, these studies do not actually demonstrate the involvement of RAL in oncogenesis or metastasis development. Again, the use of ectopically expressed proteins in the experiments conducted, do not mimic those seen in physiological situations. The resulting protein overexpression used in these methods, may lead to improper subcellular localisation and in some cases

constitutively activated endogenous counterparts. Therefore, knockout cells that can be used to show that the loss of RAL expression reverts the transformed phenotype of tumour cells are highly beneficial.

Together, the data currently available highlights the need for an appropriate model that can be utilised to study general RAL isoform function as well as any role the pathway plays in cell dysregulation and cancer progression. The generation of a panel of isoform-specific gene edited cell lines would overcome many of the problems associated with knockdown and overexpression studies and ultimately help elucidate the importance of RAL signalling downstream of oncogenic RAS.

3.1.1 Suitable RAL relevant cell lines

Since RAL primarily acts as an effector of active RAS, it seemed sensible to choose a RAS model focusing on one specific cell type to study RAL isoform biology. We also wanted to factor in RAL dependency on oncogenic RAS, particularly the most frequently mutated isoform, KRAS. With mutations in all three RAS oncogenes present in over 50% of colorectal carcinomas (Serebriiskii *et al.*, 2019), colorectal cell lines are a suitable choice when studying RAL.

RALA and RALB have also been shown to demonstrate antagonistic roles in colorectal cancer, with RALA activation required for anchorage-independent growth, whilst RALB suppression enhanced tumour survival (Martin *et al.*, 2011). This is an intriguing result that presents the idea of opposing roles for the RAL isoforms that might not just be confined to oncogenic RAS.

Table 3.1 displays the panel of colorectal cell lines chosen for consideration. The fact there is mutant and wildtype RAS present in this cell panel, is an important feature for looking at normal and disease driven RAS biology. A recent study looking at RAS and RAL contributions to cell viability via knockdown (Lee *et al.*, 2019), found a difference in terms of cell survival across this particular panel of colorectal cell lines. However, the reasons for the difference in RAL dependency between WT and mutant KRAS and even within various KRAS mutations remains unclear. Focusing on these colorectal cell lines, will provide the opportunity to discover new things about RAL function and its involvement in oncogenic RAS signalling.

Table 3.1 Mutational status of colorectal cancer cell lines considered.

Cell Line	Cell Type	KRAS Status	RAL/RAL Regulator Status	APC Status
CaCo2	Colon, adenocarcinoma	WT	RALA, RALB: WT	Q1367*
DLD-1	Colon, adenocarcinoma	G13D	RALA, RALB: WT RALGAPA2: M1043V RALGAPB: R717H RALGDS: S462I, S490I RALGPS1: A481V, A531V RALGPS2: Q385K	R727M K993N R2166*
HCT116	Colon, adenocarcinoma	G13D	RALA, RALB: WT RALGAPA1: Q1432Q RALGPS1: P420L	WT
LoVo	Colon, adenocarcinoma	G13D	RALA, RALB: WT RALGPS1: P380P	R1114* R2816Q
SW48	Colon, adenocarcinoma	WT	RALA, RALB: WT	R2714C
SW403	Colon, adenocarcinoma	G12V	RALA, RALB: WT RALGAPA2: L212I	N125K
SW620	Colon, adenocarcinoma	G12V	RALA, RALB: WT RALGAPA2: W620L	Q1338*

KRAS mutation status was obtained from Broad Institute Cancer Cell Line Encyclopaedia and verified on COSMIC. All heterozygous except SW620. Abbreviations: WT (wild type), *nonsense mutation.

3.1.2 Aims and Objectives

To screen a panel of colorectal cell lines to assess whether they would be potentially suitable for CRISPR/Cas9-mediated gene editing and subsequent study of RAL isoform biology. Assessment consisted of ease of handling relevant to gene editing protocols and endogenous RAL activity levels.

3.2 Results

3.2.1 Ease of handling

Ease of handling is an important factor for selecting cells that are amenable to CRISPR/Cas9 gene editing. This focused on two key criteria:

i) Transfection efficiency – this increases the likelihood of gene editing taking place.
ii) Single cell viability – a key step in the selection of gene edited cell clones.

Transfection efficiency

Transfection of a plasmid expressing GFP (Supplementary Supplementary Figure 7.7) using standard lipid transfection reagents was carried out to assess plasmid uptake efficiency. Although lentiviral particle production is an option for generating CRISPR/Cas9 edited cell lines, the large size of the Cas9 enzyme leads to difficulties in viral particle packaging and additional steps would be required to aid transduction, lengthening the overall procedure time. Since the LentiCRISPR plasmid also contains an antibiotic selection marker, transfection using normal lipid transfection was deemed a sufficient technique for generating isoform specific RAL knockouts. Poor GFP expression was observed in CaCo2 and SW403 cell lines (Figure 3.1). Visible GFP expression was seen in the remaining colorectal cell lines tested with HCT116 cells showing GFP fluorescence levels similar to those observed in HEK293T cells.

Amenability for single cell clonal selection

Trypsin with 0.5% EDTA cell dissociation reagent was systematically profiled across the panel of cell lines to determine the ease of single cell isolation. As well as the laboratory standard, alternative dissociation reagents were investigated and StemPro Accutase was identified as a convenient and cost-effective detachment solution, as well as being a direct substitute to trypsin, since it is made up of proteolytic and collagenolytic enzymes not derived from mammalian or bacterial products.

However, even with the use of these different reagents, some cell lines still didn't dissociate. Several of the cell lines, most notably CaCo2, remained clumpy (upwards of 5 cells) following dissociation tests, as did DLD-1 and LoVo, although to a lesser extent. The inability to grow these cells from single cell clones meant that these were automatically excluded from final selection. Partial isolation, in which a mix of single cells and cell clumps is achieved, was observed in two cell lines, HCT116 and SW48. As a result, out of all cell lines described in Table 3.1, HCT116, SW48, SW403 and SW620 were the only ones that could be reliably cultured from single cell clones. These four cell lines were therefore retained for the next phase of testing.

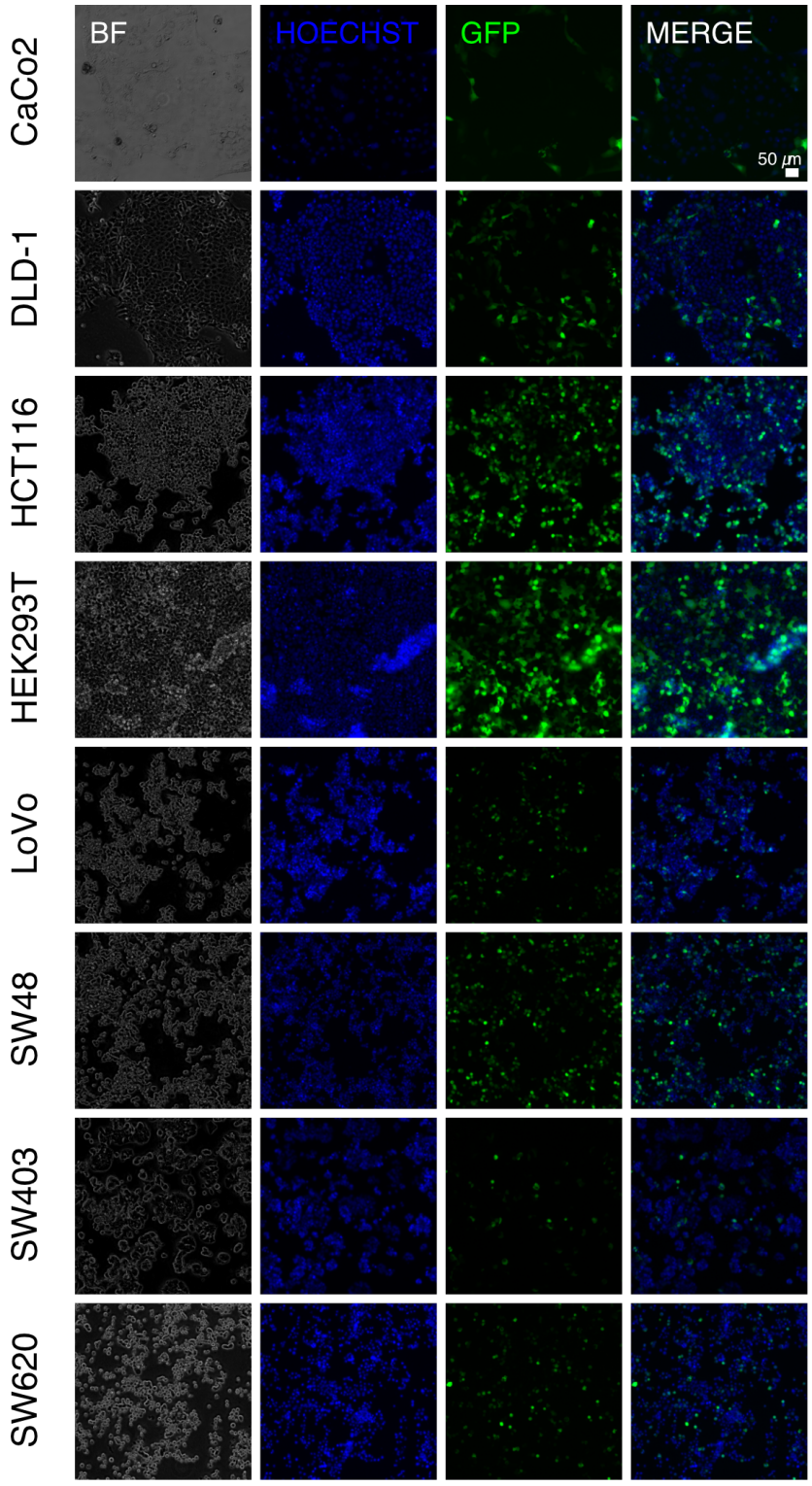


Figure 3.1 GFP expression following lipid-based transfection.
GFP expression following transfection of pEGFP-C1 using Lipofectamine LTX PLUS in a panel of colorectal cell lines plus HEK293T cells at 10X magnification. GFP fluorescence is shown in green, nuclear staining in blue, BF = bright field. Scale bar: 50 μm.

3.2.2 Activation of the RAL pathway

In order to achieve the primary aim of generating gene-edited cell models investigating isoform-specific RAL biology, it was important to select cell lines with RAL pathways that were demonstrably functional. An essential requirement for the selection of suitable cell lines was that RAL could be activated in response to upstream stimulation. An overview of the RAL activity assay is shown in Figure 3.2. Following stimulation of RAS through EGF at various timepoints, cells were lysed and incubated with a GST fusion protein containing the RAL binding domain of RAL effector RALBP1 (RLIP76) (Wolthuis *et al.*, 1998a). Since this protein only binds to the GTP-bound active form of RAL, both isoforms were isolated and the activity of each determined by western blotting. Probing for phospho-ERK, a known RAS effector in the MAPK pathway, was also carried out to confirm positive RAS activation by EGF treatment.

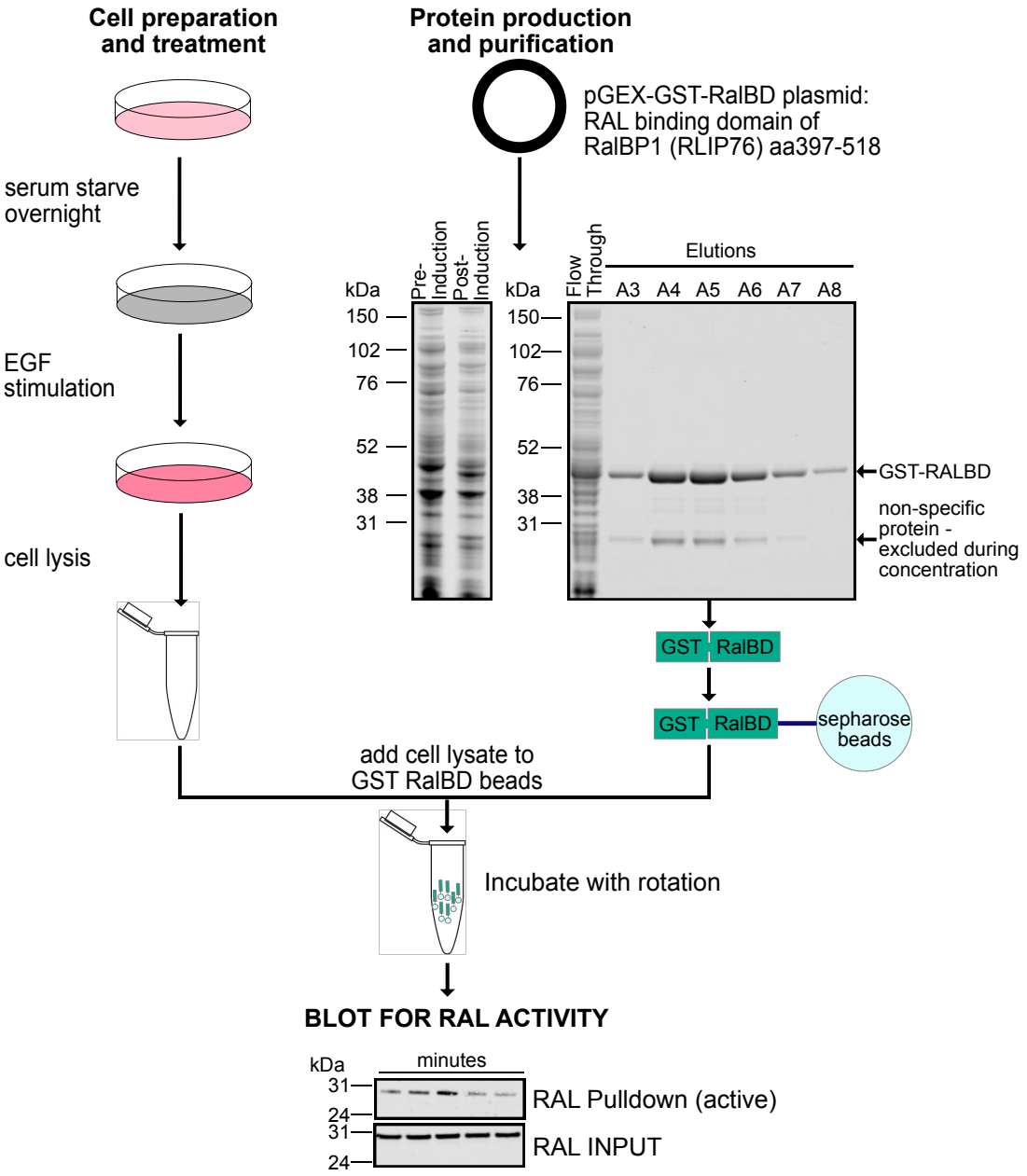


Figure 3.2 An overview of the RAL activity assay.
Following stimulation of the RALGEF/RAL pathway with EGF, cell lysate is incubated with GST-tagged RALBD (RAL binding domain of RAL effector RALBP1) and sepharose beads. The beads are subsequently washed and resolved by SDS-PAGE before probing for RALA and RALB.

Overall, all cell lines except one showed an increase in RALA and RALB activity (Figure 3.3 & Figure 3.4). In general, peak RAL activity was observed following 1-2 minutes of EGF treatment. No obvious difference was seen between the activity of one isoform compared to the other in any cell line. The total protein expression for RALA and RALB remained the same throughout the time course in all cell lines.

Cell line HCT116 displayed the fastest but also the most short-lived activation of both isoforms following only 1-minute EGF treatment, before returning back to unstimulated levels (Figure 3.3). A clear and early activation of RALA and RALB was also observed in the SW48, peaking at 2 minutes and lasting a total of 5-10 minutes (Figure 3.3). SW403 cells also showed an early activation of RAL with just 1-minute EGF treatment (Figure 3.4). Unlike the HCT116 cell line, RAL activity was prolonged until 5 minutes stimulation. Only SW620 showed no discernible increase in active RAL expression with EGF treatment (Figure 3.4). Instead, a constant level of active protein expression can be seen at all time points including untreated.

HCT116 and SW48 cell lines displayed a clear and steady increase in phospho-ERK. This activation was later than that of RAL and started from around 5 minutes. A clear increase in phospho-ERK levels could also be seen in the SW403 line, peaking around 5-10 minutes after EGF treatment initiation. Only SW620 displayed a sharp increase of phospho-ERK at 1 minute, before returning quickly back to the levels seen in serum-free medium.

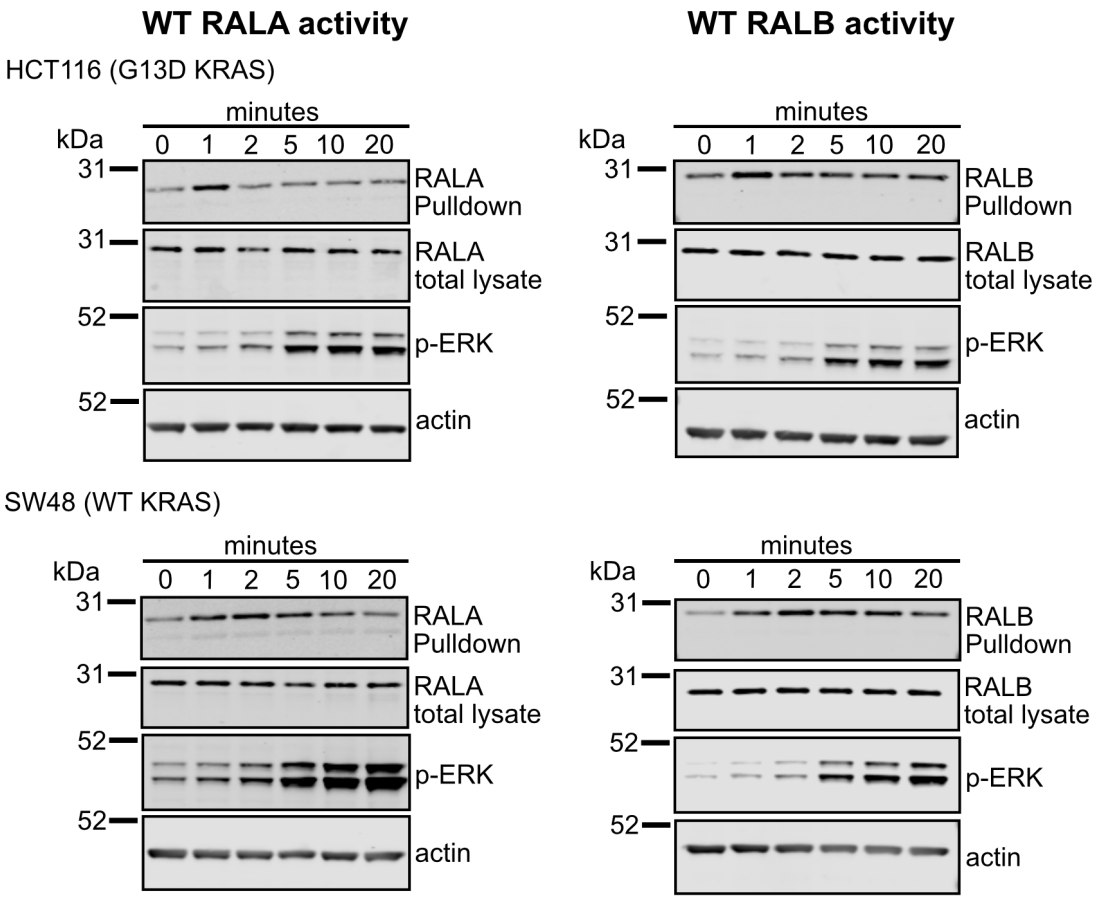


Figure 3.3 RAL activity in parental HCT116 and SW48 cell lines.
Following treatment with 50 ng/mL EGF at various time points (0-20 minutes), HCT116 cells displayed a clear increase in activity of both isoforms at 1 minute before quickly returning back to untreated levels. A clear and early activation of RALA and RALB was also observed in SW48 parental cells with peak activation around 2 minutes. RAL activity in the SW48 was longer lived, lasting 5-10 minutes in total. Phospho-ERK was also probed to confirm successful activation of RAS. n=1.

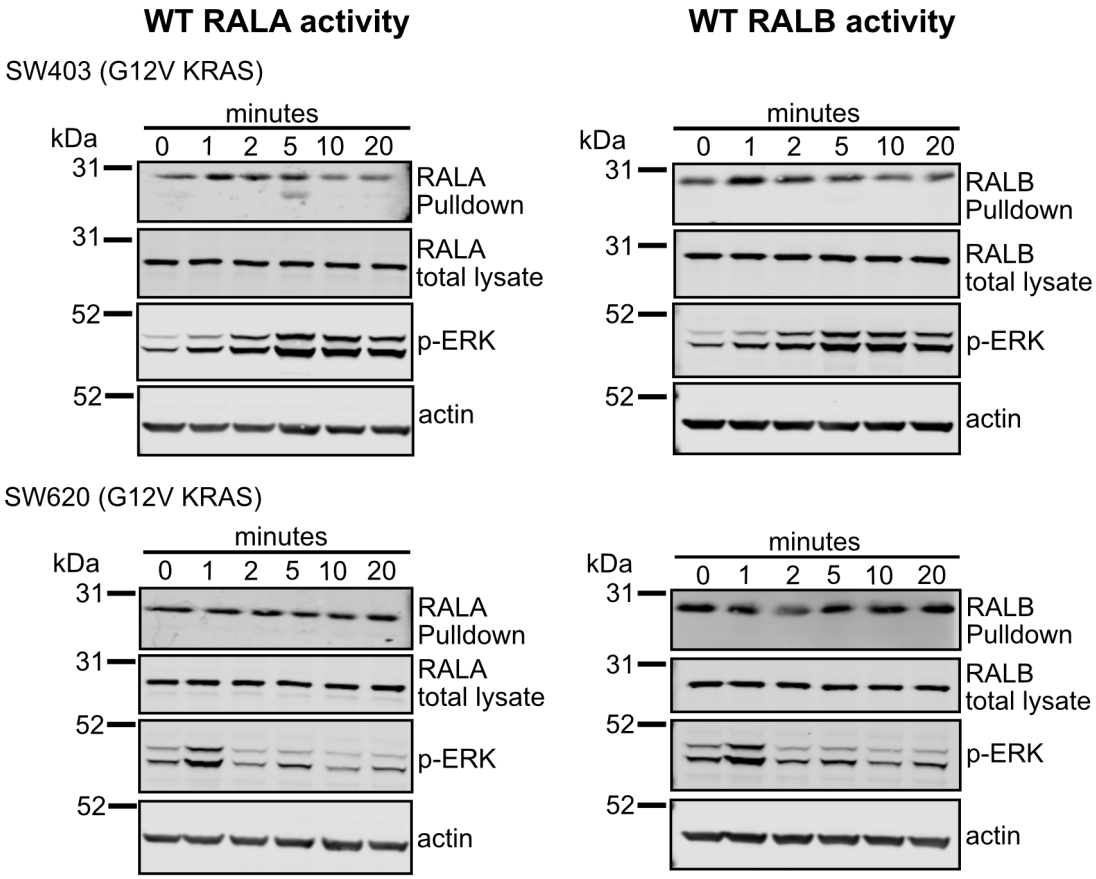


Figure 3.4 RAL activity in parental SW403 and SW620 cell lines.
Following treatment with 50 ng/mL EGF at various time points (0-20 minutes), SW403 cells displayed a clear increase in activity of both isoforms at 1 minute that lasted up to 5 minutes. SW620 parental did not show any obvious activation of either isoform at any time point. Instead, steady expression levels could be seen throughout. Phospho-ERK was also probed to confirm successful activation of RAS. SW403 showed an increase in phospho-ERK over time, peaking around 5-10 minutes. SW620 displayed an increase in phospho-ERK at 1 minute before returning back to baseline levels. n=1.

Table 3.2 Desirable features for cell line selection.

Name	KRAS Mutation	RAL Regulator mutations	Easily Transfected	Single cell isolation	RAL activation
CaCo2	NO	NO	NO	NO	-
DLD-1	YES	YES	YES	NO	-
HCT116	YES	NO	YES	SOME	YES
LoVo	YES	YES	YES	NO	-
SW48	NO	NO	YES	SOME	YES
SW403	YES	YES	NO	YES	YES
SW620	YES	YES	YES	YES	NO

Table 3.2 summarises the cell lines under consideration and their desirable feature scores to aid final cell selection. Four different cell lines were selected to undergo gene editing in an effort to reduce the dangers associated with cell line and clonal variability. The following ticked most of the boxes, HCT116, SW48, SW403 and SW620.

3.3 Discussion

In this chapter, an initial pre-screen of cell lines was performed to assess whether they would be potentially suitable for CRISPR/Cas9-mediated gene editing and subsequent study of RAL biology.

Generating a panel of gene-edited cell lines from a particular cancer type provides more confidence into any insights that were generated. This is because signalling is often context-dependent and responses from a single cell line may not be representative of the general biology being investigated. Colorectal cancer was chosen as the model system because of its high relevance to RAS (~50% of colorectal cancer patients harbour mutations in KRAS (Serebriiskii *et al.*, 2019)). A recent synthetic lethality study using the cell lines in this panel showed knockdown of RAL holds greater toxicity in mutant KRAS cell lines compared to WT KRAS (Lee *et al.*, 2019). Here, KRAS oncoeffector nodes were determined by which effector knockdowns most closely phenocopied that of KRAS loss. RAL was identified as a key node in

the mutant cell lines, with ~50% decrease in cell viability observed in HCT116, and up to 25% decrease in viability in SW403 and SW620 cell lines. This gives confidence that RAL biology is relevant to these cells. However, the distinct requirements for RAL across the cell lines also show that there are still important aspects of RAL contributions to oncogenesis and RAS signalling that remain to be discovered.

Arguably the most important factors to consider were ease of single cell isolation and transfection efficiency for the generation of isogenic cell lines. These were the initial cut-off points for cell selection. Once the panel was systematically profiled for these criteria and unsuitable cells excluded, the ability to display high/clear RAL activation was also deemed highly desirable. Although expected, it was difficult to acquire cell lines that satisfied every one of these criteria. Transfection using normal lipid-based reagents was variable across the whole panel of cell lines. Fortunately, with the inclusion of an antibiotic selection marker, even minimal transfection efficiency is more than sufficient for generating knockout cell lines. Very poor transfection efficiency combined with the inability to isolate single cells resulted in the exclusion of the cell lines CaCo2, DLD-1 and LoVo before the next step of testing.

SW48 did not score highly for all desirable features, but since it was important to include a cell line that was WT for KRAS and they were easier to handle than CaCo2 cells, they were chosen for further investigation. The addition of a WT KRAS cell line gives us the option of exploring RAS-driven and RAS-independent RAL biology. SW620 is the only line to contain a homozygous KRAS mutation. For this reason, as well as being the easiest to isolate single cells, it was also included in the final cell selection. Mutations in KRAS are recorded as being typically heterozygous, so the presence of a homozygous KRAS cell line might enhance RAL activity and therefore help clarify RAL mechanisms. It may also address any contribution RAL plays as a result of RAS mutant allele-specific imbalance (MASI), in which the mutant allele becomes dominant as a result of WT deletion or mutant copy number gain (Krasinskas *et al.*, 2014; Krasinskas *et al.*, 2013).

Utilising cell lines that have an easily activated RAL pathway or high RAL activity in general are optimal for isoform function analysis. High activity places a certain degree of dependency on a cell line, with changes in that pathway potentially leading to alterations in cell function and viability.

EGF is a potent growth factor when it comes to activating RAS and the RAF pathway, with negligible RAS effector activation seen in serum starved cells regardless of RAS status (Hood *et al.*, 2019). Indeed, stimulation of various growth factor receptors including EGF, results in rapid activation of RAL and a dominant negative RAS completely inhibits RAL activation induced by EGF (Wolthuis *et al.*, 1998b).

Although the activity assays in section 3.2.2 are n=1, the increase in both RAL isoforms at multiple time points gives me confidence that RAL was being engaged and these cell lines are responsive. Both the SW48 and the SW403 cell lines have active RAL that is sustained for several minutes. There was no clear increase in activity observed in the SW620 cell line. This was repeated with my own beads as well as a commercial kit (data not shown) and it failed in both cases to see a response. Instead, the signal for protein expression at every time point is as strong as the peak activity recorded in the other cell lines. This could suggest the SW620 cell line has an active RAL pathway that remains high, regardless of stimulation or is faster/slower than the time points tested. The RALGEFs RALGPS1A, RALGPS1B and RALGPS2 are thought to activate RAL downstream from growth factor receptors independently of RAS activation (de Bruyn *et al.*, 2000) and therefore the active RAL observed may in part be due to alternate growth factor signalling pathways.

The duration of RAL activity in all the cell lines tested is relatively short in comparison to RAS activity and other RAS effectors. Transient activation of RAS has been observed for as long as 10 minutes in both HeLa and MEF cells following EGF stimulation (Hennig *et al.*, 2016). Moreover, activity of RAS effector ERK was detected for 30 minutes in the same HeLa cells. This is in comparison to the average 1-2-minute activity of both RALA and RALB seen here, which agrees with the RAL activity duration observed in A14, Rat1 and CCL39 cells following growth factor stimulation (Wolthuis *et al.*, 1998b).

Interestingly, the SW620 cell line is also the only one to not have a gradual increase in phospho-ERK expression that is subsequently sustained for

several minutes. Rapid, transient activation of ERK by EGF has been reported in PC12 cells (Traverse *et al.*, 1992). It is proposed that a mechanism exists for rapidly terminating EGF action on the MAPK cascade, even if the growth factor is still present, and that this rapid inactivation is likely to take place near or at the level of the EGF receptor. This could also explain why the RAL activity in the SW620 was not observable at the time points tested, even at 1 minute. Surprisingly few studies have recorded RAL activity in this way and in these particular cell lines. Usually as a result of RAL regulator alteration, activity has mainly focused on the overall activation state of individual RAL isoforms. Increased RALA activation after stimulation with 10% FBS for 5 minutes was observed in oral squamous cell carcinoma cell lines following knockdown of RALGAP (Gao *et al.*, 2019). RALA activity has been shown to significantly decrease in HeLa cells following CCCP treatment to induce mitochondrial depolarisation (Pollock *et al.*, 2019). Here, the RAL-GTP levels were determined over a time course lasting 80 minutes, with little insight into the activation status of both isoforms within the initial 30 minutes. Transfection of Cos7 cells with either WT or active RAL recorded approximately 8% of WT RAL was GTP bound, whilst over 85% of the active version contained GTP (Wolthuis *et al.*, 1998a). Although the pulldown method used in all these studies was similar to the one used in section 3.2.2, none of them elucidate the initiation or duration of endogenous RAL activity in an appropriate RAS model.

It is also unclear whether the additional mutations in key RAL regulators will interfere with data analysis and interpretation of normal RAL signalling. According to the COSMIC database, the RALGAPA2 missense mutation (W620L) present in the SW620 has a high pathogenic score of 0.98 (out of 1), as does the mutated RALGPS1 (P420L) in HCT116 cells (score 0.95) (Rogers *et al.*, 2018). However, none of the resulting mutated proteins present are classed as tier 1 genes by COSMIC. This means no evidence of activity that may drive cancer has been provided and therefore these mutations are unlikely to impact any oncogenic role that RAL might play in the context of RAS. Nevertheless, mutational statuses of all RAL GEFs and GAPs will be taken into consideration when analysing activity data.

RAL itself has also been shown to be a critical regulator of Wnt signalling as well as the maintenance of the intestinal stem cell population in mice (Johansson *et al.*, 2019). Genetic deletion of either isoform resulted in reduced stem cell function and hypersensitivity to Wnt inhibition. Whilst rapid crypt death was observed following ablation of both RALA and RALB. This offers up the idea of potential crosstalk between the RAS and Wnt signalling pathways through RAL activity. Therefore, it should be highlighted that dysregulation of the Wnt/β-catenin pathway plays a major role in colorectal carcinogenesis (Fodde, 2002). The presence of a mutation in the APC gene occurs more frequently than that of KRAS in sporadic colorectal cancer (Gerecke *et al.*, 2013). Indeed, many colorectal cancer cell lines available mirror this stance as they contain various mutations including the common R1450 nonsense mutation (Cheadle *et al.*, 2002). This alteration in APCs mutation cluster region results in the truncated protein and subsequent Wnt signalling dysregulation. None of the cell lines included in this panel harbour mutations in this particular region and the missense mutations present in the final cell lines chosen, score low to 0 on the mutation “hotspot” scale (Broad institute CCLE, APC gene search). This could suggest minimal interference from potential aberrant APC signalling which will assist in any future analysis of the role RAL plays in this pathway.

Chapter 4: Editing RAL isoforms using the CRISPR/Cas9 System

4.1 Introduction

At the beginning of my project, there were no knockout or specific mutation knock-in human cancer cell models targeting RAL available to study endogenous isoform functions. As a result, our knowledge of RAL isoform specific biology is limited. With particular emphasis on RALA being the more dominant isoform, RAL isoform function has centred on specific subcellular interactions with little expansion into RAL isoform biology as a whole. Therefore, editing the gene locus and generating isoform specific knockout and RAL^{G23V} knock-in cell lines was desirable. This cannot be achieved through normal transfection methods and instead a programmable nuclease system is required.

The use of programmable nucleases to edit the genome is a rapidly evolving field that enables the disruption of target genes in both an *in vitro* and *in vivo* context. Whilst nucleases such as Zinc-finger nucleases (ZFNs) and Transcription activator-like effector nucleases (TALENs) have been shown to be and still are an effective tool for genomic manipulation (Urnov *et al.*, 2010; Bogdanove and Voytas, 2011), they are limited by the need to generate a specific protein for each DNA target site. Thus, due to its simplicity and efficiency in triggering site-specific DNA cleavage, it is the CRISPR/Cas9 system that has become the preferred gene editing technique.

Whilst several CRISPR/Cas9 systems have been characterised and developed for research purposes, it is the SpCas9 derived from *Streptococcus pyogenes* that was the first to be used successfully in humans and is the most commonly used Cas9 nuclease for gene editing (Jinek *et al.*, 2013).

This has been made possible through human codon optimisation of the Cas9 nuclease as well as design of relevant and efficient guide RNA (gRNA) components. This includes a CRISPR RNA, which contains the gRNA, fused to a trans-activating CRISPR RNA through a linker loop region to form the final single gRNA (Kim and Kim, 2014). A protospacer adjacent motif (PAM) is a small base pair sequence directly upstream of the Cas9 nuclease binding site

on the target DNA strand. Following binding to the target sequence, Cas9 will cause a double stranded break in the DNA, three base pairs downstream of the PAM sequence. Subsequent repair of these breaks occurs through one of two pathways; non-homologous end joining (NHEJ), or homology directed repair (HDR). The former, NHEJ, is highly error prone and usually results in the introduction of an indel (insertion/deletion) mutation and a loss in function of the target gene. Conversely, HDR requires the addition of a homologous sequence template to allow precise repair and the option of knock-in gene editing (Chapman, Taylor and Boulton, 2012).

Excluding CRISPR screens, RAL isoforms have only been targeted using the CRISPR/Cas9 system, with the intention for further experimental study, in HEK293T cells (Johansson *et al.*, 2019). Although these knockouts were used to infer the involvement of RAL in Wnt signalling, it is not a suitable cell model to study RAL in the context of RAS biology. Therefore, a CRISPR/Cas9 strategy to knockout separate RAL isoforms across the cell lines selected in chapter 3 needed to be established. This would also enable the generation of the tools required to insert a glycine to valine mutation in RAL at amino acid position 23 via a “knock-in” resulting in a constitutively activated RAL protein. Establishing a panel of RAL knockout and RAL^{G23V} cell lines will allow us to analyse the role of RAL isoforms in different cellular contexts.

4.1.1 Aims and Objectives

The aim of this chapter was to generate isoform specific RAL knockout lines in a panel of colorectal cancer cell lines, as well as optimise the tools needed to establish RAL^{G23V} knock-in cell lines.

The following steps were taken to achieve this:

- Validate the targeting efficiency of four different gRNAs per RAL isoform in HEK293T cells.
- Optimise the double gRNA approach in HEK293T cells before applying to target cell lines.
- Design and generate a construct containing a HDR DNA template with a G23V mutation for knock-in cell lines.

4.2 Knockout cell lines

All target cell lines, initially underwent single cell isolation and expansion to generate parental clones to reduce clonal variability. These parental clones then underwent gene editing using CRISPR/Cas9.

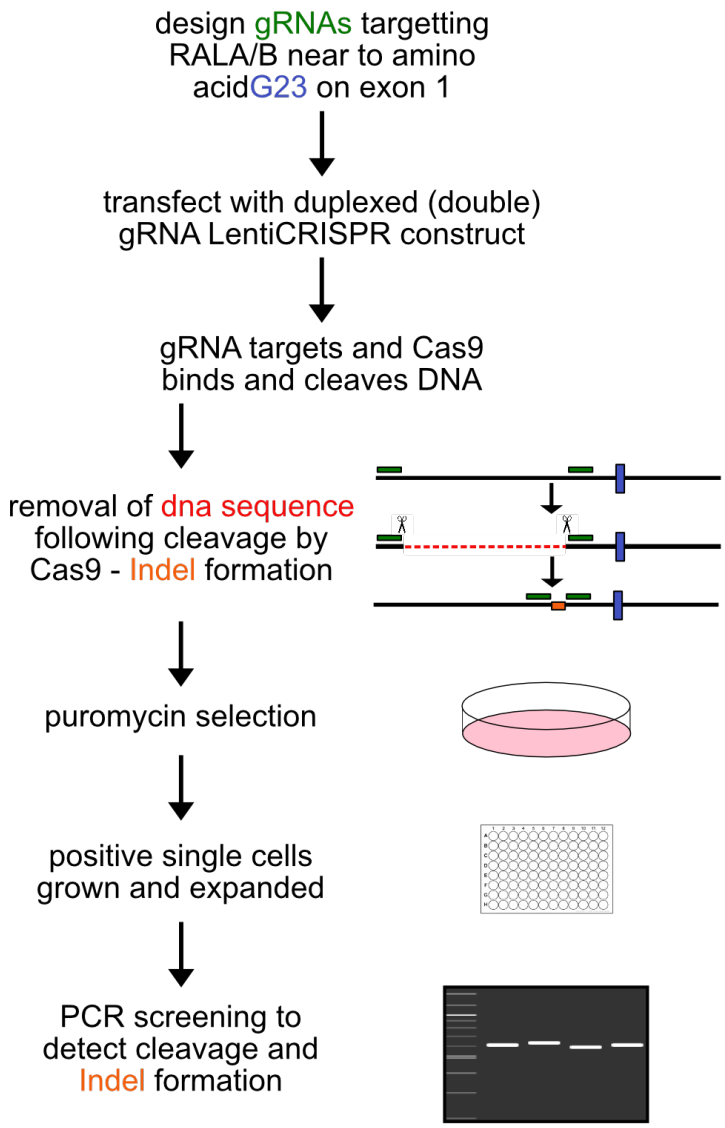
4.2.1 Double gRNA approach

Using a single guide to target a gene of interest and generate knockout cells often results in small indels that are difficult to detect on a standard agarose gel and require assays such as the T7E1 to assess editing efficiency. By using a paired guide system that targets the same gene hundreds of bases apart, larger deletions can be detected on an agarose gel following amplification of the targeted region (Figure 4.1A).

4.2.2 Guide RNA design

All CRISPR/Cas9 components target the Exon 2 of each RAL gene, which is where the protein coding region initiates. Therefore, this region is subsequently referred to as Exon 1 from this point on. Guide RNAs (gRNAs) targeting Exon 1 of RALA/B (Figure 4.1B) were designed using the Synthego online tool https://design.synthego.com/#/validate. Guides were chosen for their low off-target effects but also for their proximity to the start codon in both RAL isoforms as well as amino acid G23. Exon 1 gRNAs were chosen as indel (insertion/deletion) mutations early on in a gene, as they are much more likely to lead to gene disruption.

A)



B)

RALA	Sequence	Direction	Location
gRNA #1	ACTCAGTATCCGAACAAAAT	antisense	Intergenic
gRNA #2	GATTCACTCGTATGAAGTAA	antisense	Intergenic
gRNA #3	TTTACACAAAGTCATCATGG	sense	Exon 1
gRNA #4	AAAGTCATCATGGTGGGCAG	sense	Exon 1
RALB			
gRNA #1	TGTCATCAGCAGCTCTTCAG	antisense	Exon 1
gRNA #2	GTCATCAGCAGCTCTTCAGT	antisense	Exon 1
gRNA #3	AAGGTGATCATGGTTGGCAG	sense	Exon 1
gRNA #4	GTGATCATGGTTGGCAGCGG	sense	Exon 1

Figure 4.1 An overview of the generation of isoform specific RAL knockout cell lines using the CRISPR/Cas9 system.
A) Outline of steps used to create isoform specific knockout cell lines using a LentiCRISPR construct containing two gRNAs and the Cas9 enzyme. **B)** gRNA sequences targeting Exon 1 of RALA/B near G23 amino acid were designed using the Synthego online tool.

4.2.3 Design and construction of LentiCRISPR constructs

For generation of knockout cell lines using the CRISPR/Cas9 system, delivery of the essential gRNA and Cas9 benefits from the use of a single plasmid. In this instance, all knockout cell lines utilised a modified form of the LentiCRISPR V2 plasmid https://addgene.org/52961/ (Addgene Boston USA) which was a kind gift from Dr Nicholas Harper (University of Liverpool, UK). This plasmid has had the minimal EFS promoter replaced with a Spleen Focus-Forming Virus (SFFV) promoter in order to enhance Cas9 expression. This plasmid contains an enhanced Streptococcus pyrogenes Cas9 (eSpCas9) with a C-terminal Strep tag II motif fused to the PacI gene to confer Puromycin resistance (Supplementary Supplementary Figure 7.6). LentiCRISPR-SFFV was digested using *Bsm*BI for 3h at 55°C. Digested plasmids were analysed on a 0.75% agarose gel to ensure removal of the non-essential "stuffer" DNA, to allow insertion of the gRNAs and purified using the Wizard™ SV Gel and PCR clean up system. Each gRNA was ligated into individual LentiCRISPR-SFFV vectors. This resulted in 8 plasmids in total, 4 with gRNAs targeting RALA and 4 targeting RALB.

4.2.4 Optimising genotyping primers

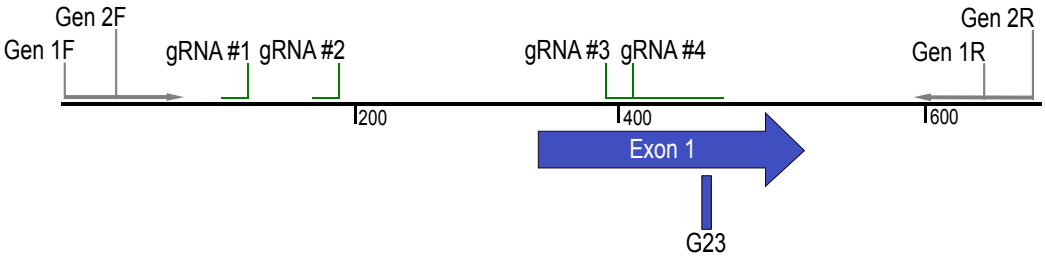
To screen for successful Cas9 cleavage, primers were designed outside the gRNAs in the Exon 1 region (Figure 4.2A & B). This should result in a full-length PCR product in unedited cells and a shorter product in edited cells due to the loss of DNA between the target gRNAs. Amplification using all primer combinations was successful and produced a PCR product of ~600 bp on WT genomic DNA from HEK923Ts (Figure 4.2C). Genotyping primer combinations, Gen 1F + Gen1R for both RALA and RALB were chosen for future knockout screens.

A)

Genotyping Primers RALA
Gen 1F 5'-GCT GAT ACC CCA TTA TTG TCA GC-3'
Gen 2F 5'-GTC AGC ATG CAG TTG TAC TGG G-3'
Gen 1R 5'-AAT CAA GTT ATG TAC AGA TGG CTC-3'
Gen 2R 5'-CCA TAA GTT CCT GAT TTG TAA AAT C-3'

Genotyping Primers RALB
Gen 1F 5'-GTG ACT TCC AGA GTT GTT GG-3'
Gen 2F 5'-GAG GGC TCG CAT GTC TGC-3'
Gen 1R 5'-CGA TGC AAG CAG AGA TCT CTC-3'
Gen 2R 5'-TAA CAA TGG AAA CGA TGC AAG-3'

B)



C)

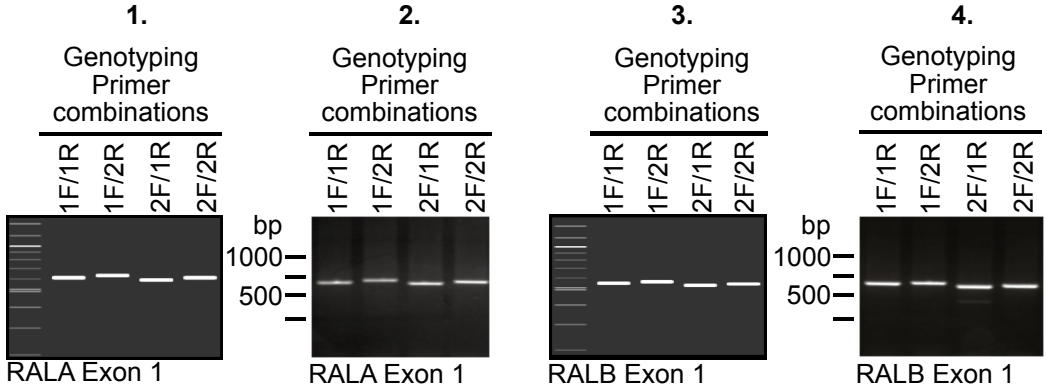


Figure 4.2 Genotyping primers design and optimisation for generation of isoform specific RAL knockout cells.
A) Genotyping primer pairs for RALA and RALB isoforms, forward and reverse. **B)** Example schematic of the position of genotyping primers in relation to RAL gRNAs and Exon 1. **C)** 1+3 are gel mock-ups of predicted PCR amplification of RALA and RALB Exon 1. 2+4 are confirmed amplification of unedited (UT) RALA and RALB Exon 1 (PCR product ~600 bp) using designed genotyping primers.

4.2.5 Single gRNA LentiCRISPR plasmids modify isoform specific RAL in HEK293Ts.

To determine gRNA targeting and Cas9 cleavage efficiency, each single gRNA LentiCRISPR plasmid was validated in the easily transfectable cell line HEK293T. HEK293T cells were transfected with a paired combination of the four gRNAs specific to each isoform.

Following amplification of Exon 1 using the genotyping primers specified in Figure 4.2A, untransfected (UT) cells produced a band of ~600 bp (Figure 4.3A). Co-transfection of all gRNA combinations produced a smaller product for both RALA and RALB in around half of the cell population (Figure 4.3A) indicating DNA has been excised from Exon 1.

Confirmation of Cas9 cleavage at the processing sites of each gRNA was verified by comparing the cut (lower band) sequence of DNA against the upper band of the unmodified wild type (WT) RALA/B (Figure 4.3B).

Although all gRNAs resulted in cleavage and editing for both RAL isoforms, gRNA#2 and gRNA#4 for both RALA and RALB were chosen for future gene editing of selected cell lines, due to this gRNA combination having the highest population of edited cells.

4.2.6 Duplexed gRNA LentiCRISPR plasmids modify RALA/B in HEK293T cells.

Once each single gRNA construct had been validated in HEK293T cells, I modified each single LentiCRISPR containing the isoform specific gRNA#2 to also contain gRNA#4 and created a duplexed guide construct (Figure 4.4A). This allows transfection of a single plasmid that will express both gRNAs required to knockout its target RAL isoform. Three separate duplexed guide constructs were transfected into HEK293T cells alongside single guide transfections acting as a positive control for gene editing. DNA was isolated and verified by PCR as previously described. All duplexed constructs targeting RALA resulted in cleavage of RALA Exon 1 (Figure 4.4B), whilst two out of three duplexed constructs targeting RALB also produced a smaller PCR product indictive of gene editing (Figure 4.4B).

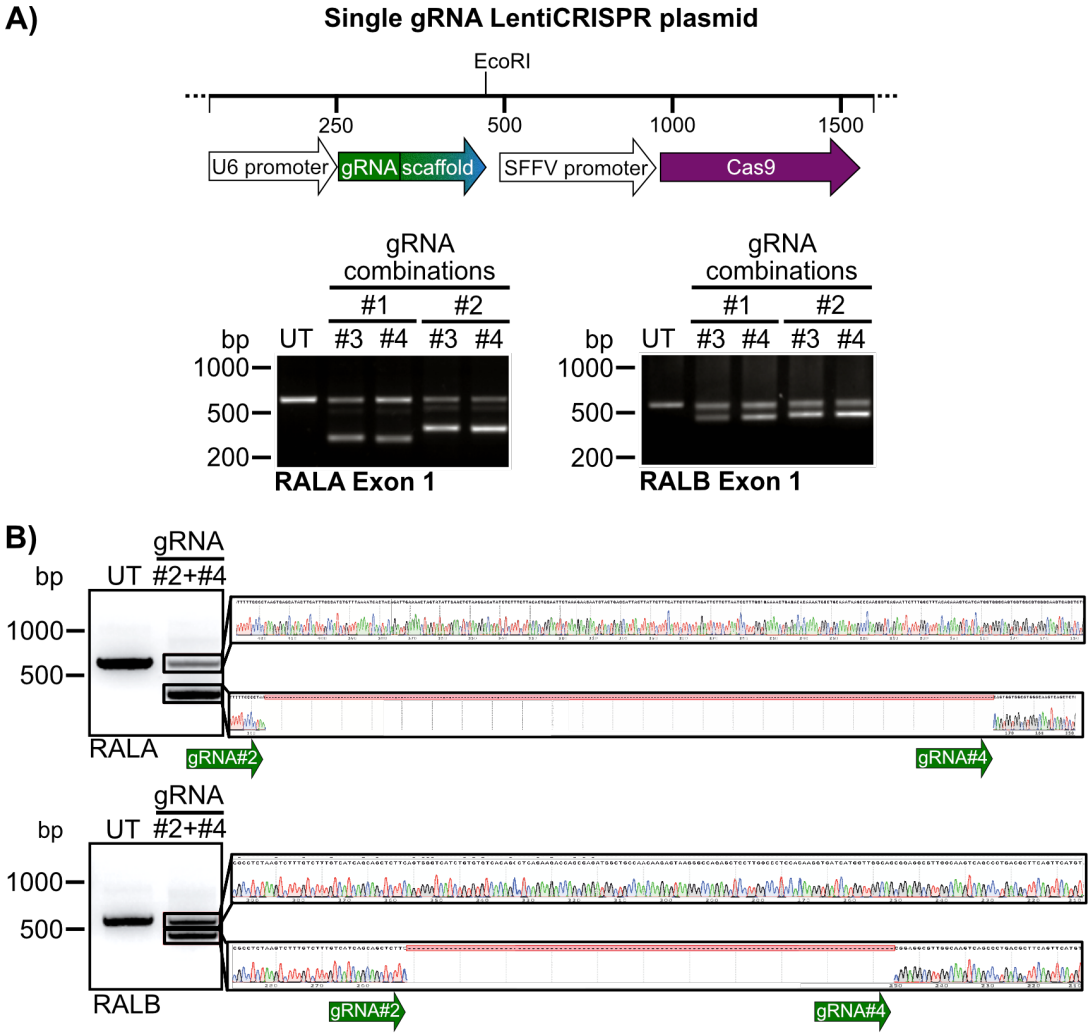


Figure 4.3 All gRNA LentiCRISPR plasmids successfully cut target DNA in both RALA and RALB.
A) Schematic of a single gRNA LentiCRISPR plasmid with PCR amplification of RALA and RALB Exon 1 in untransfected (UT) and two single gRNAs transfected HEK293T cells. **B)** Bands from the co-transfection of gRNAs #2 and #4 were extracted and sent for sequencing. Upper band represents the wild type Exon 1 of RALA/B (unmodified), whilst the lower band indicates cleavage by Cas9 nuclease.

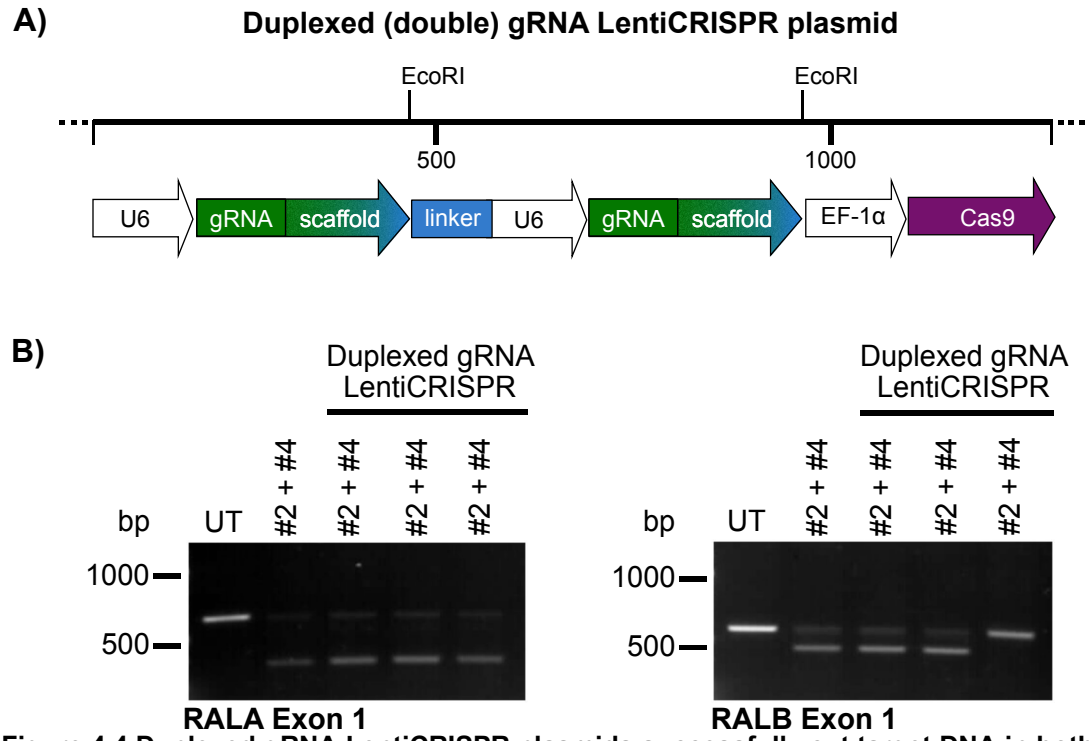


Figure 4.4 Duplexed gRNA LentiCRISPR plasmids successfully cut target DNA in both RALA and RALB.
A) Schematic of a duplexed gRNA LentiCRISPR plasmid. **B)** PCR amplification of RALA and RALB Exon 1 in untransfected (UT), 1X separate gRNA transfection and 3X duplexed gRNA plasmid in HEK293T cells.

4.2.7 Duplexed gRNA LentiCRISPR constructs modify RALA/B in target cell lines.

As mentioned previously, all target cell lines underwent single cell clonal selection to produce parental (PAR) cell lines. This was carried out to reduce clonal variability and any gene-editing was conducted on these parental cell lines.

For HCT116, SW403 and SW620, three separate parental cell lines were isolated and expanded (PAR1, PAR2 and PAR3). Only two parental lines were obtained for SW48 (PAR2 and PAR3).

Following transfection of the duplexed gRNA LentiCRISPR-RALA/B constructs into every target parental cell line, cells underwent one round of puromycin treatment to select for positive editing. The remaining cells were left to grow and then transferred into 96-well plates at a seeding density of 0.5 cells per well to obtain at least one plate per parental clone for each RAL isoform. For SW48, single cells could only be obtained from a single parental.

Seeded plates were checked by eye for single cell colonies and a maximum of five colonies was picked from each plate and subsequently expanded for DNA extraction and PCR screening.

Out of the picked SW620 clones transfected with gRNAs for RALA, four out of five (5th clone not shown) displayed a complete lower band shift in comparison to the untransfected (UT) (Figure 4.5A). No band shift was seen in the same genomic DNA screened with primers specific for RALB (Figure 4.5A). Four out of five SW620 clones transfected with gRNAs for RALB produced a smaller PCR product compared to UT (Figure 4.5A). Again, a band shift was not detected when using RALA specific primers. The second clone (BKO2) shows an incomplete band shift, with both the WT and edited band size products visible following PCR amplification. This cell line could be considered as a heterozygous population. An additional band was detected at ~400 bp in clone BKO3 as well as the predicted PCR product.

All of the SW403 clones transfected with the duplexed gRNA plasmid for RALA, displayed a complete lower band shift in comparison to the UT (Figure 4.6A). No band shift was seen in the same genomic DNA screened with primers specific for RALB. An additional band higher up than the predicted PCR product was also observed in clones AKO2 and AKO3 (Figure 4.6A). However, this was not the same size as the WT. Likewise, all SW403 clones transfected with gRNAs for RALB produced a smaller PCR product compared to UT (Figure 4.6A) with additional larger bands (~700 bp) present in clones BKO1 and BKO2. Again, a band shift was not detected when using RALA specific primers.

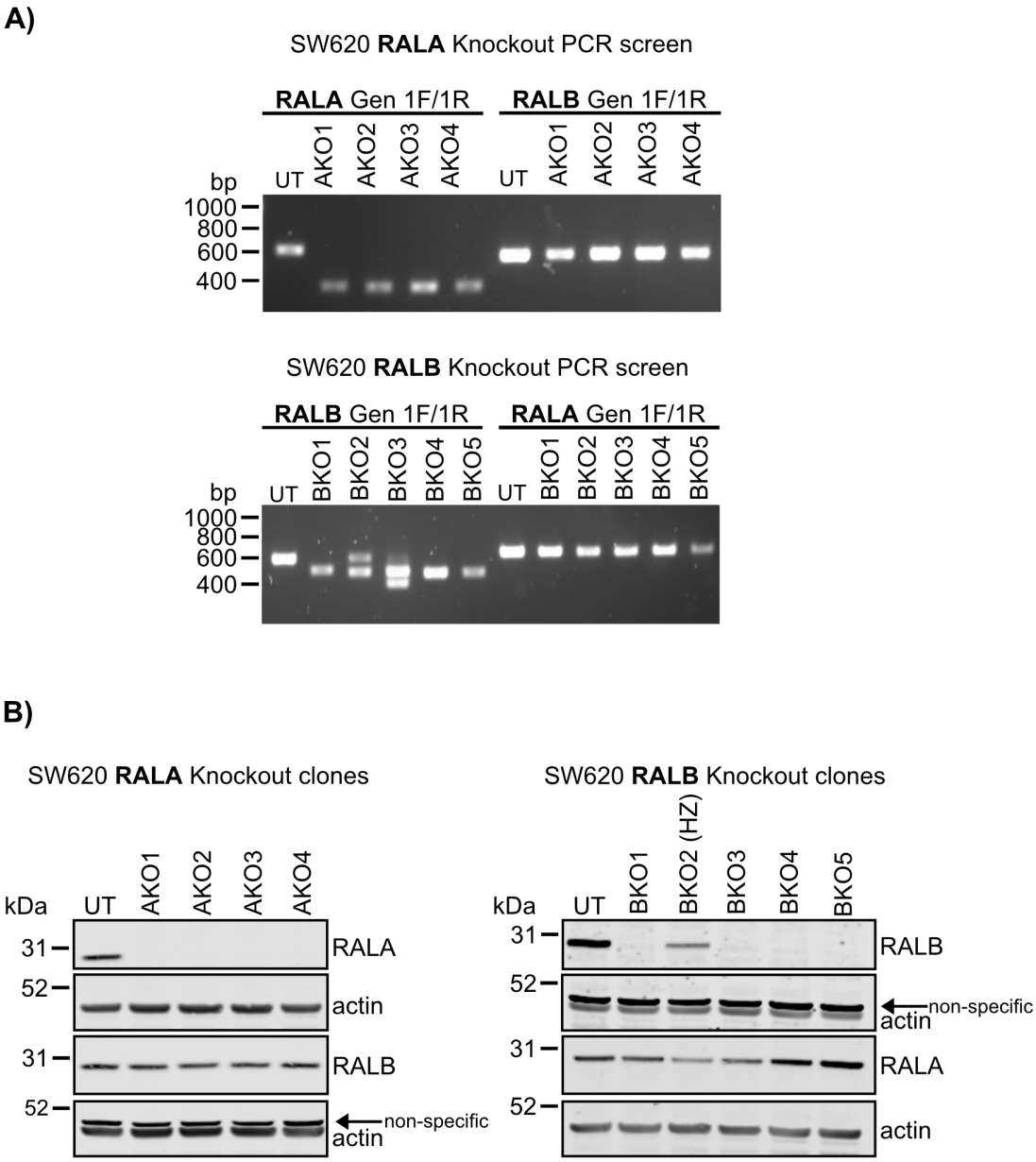
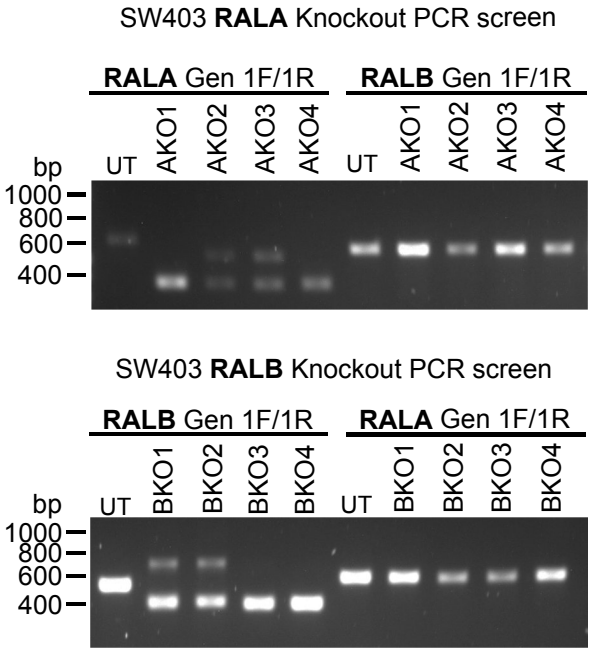


Figure 4.5 SW620 PCR screen and RALA/B knockout cell lines.
(A) Following transfection with the duplexed gRNA LentiCRISPR plasmid and subsequent antibiotic selection, four out of five SW620 clones transfected with gRNAs for RALA, displayed a complete lower band shift in comparison to the untransfected (UT) (5th clone not shown). No band shift was seen in the same genomic DNA screened with primers specific for RALB. Four out of five SW620 clones transfected with gRNAs for RALB produced a smaller PCR product compared to UT with BKO2 presenting as a heterozygous population. Again, a band shift was not detected when using RALA specific primers. **(B)** Isoform specific protein expression was absent in all positive clones identified by PCR. The total RALB expression remained unaffected in the RAL KO clones, whilst BKO2 and BKO3 had noticeably reduced RALA expression. n=1

A)



B)

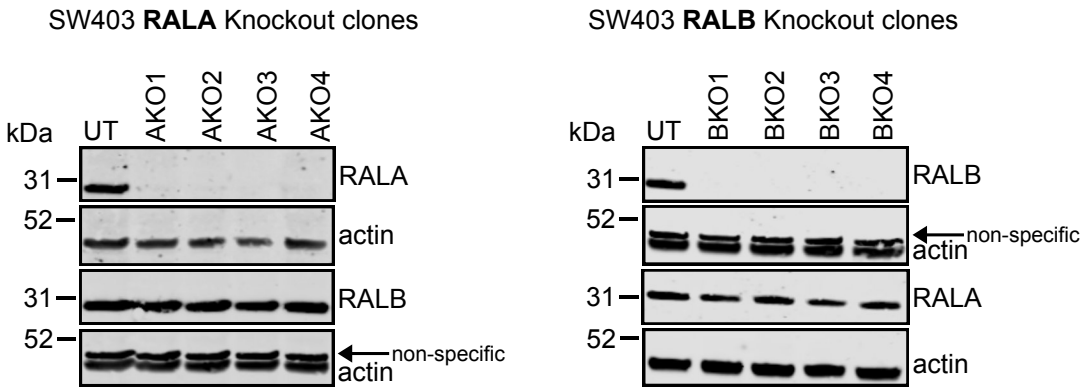


Figure 4.6 SW403 PCR screen and RALA/B knockout cell lines.
(A) All SW403 clones transfected with gRNAs for RALA, displayed a complete lower band shift in comparison to the untransfected (UT). No band shift was seen in the same genomic DNA screened with primers specific for RALB. All SW403 clones transfected with gRNAs for RALB produced a smaller PCR product compared to UT. Again, a band shift was not detected when using RALA specific primers. **(B)** Isoform specific protein expression was absent in all positive clones identified by PCR, with protein expression of the non-target isoform unaffected. n=1

4.2.8 Absence of RALA/B protein in selected cell lines.

Given the results of the PCR, isoform specific RAL protein levels were expected to be reduced or completely absent.

Analysis by western blot revealed an absence in protein production of the relevant RAL isoform in all SW620 clones (Figure 4.5B). A noticeable reduction but not complete absence of RALB was observed in clone BKO2 (Figure 4.5B). This is consistent with results from the PCR data seen in Figure 4.5A. Due to the presence of extra PCR products in BKO2 and BKO3 (Figure 4.5A), clones AKO3, AKO4 and BKO1, BKO4 were selected as the final knockout clones for future experiments. For ease, these will now be referred to as clones AKO1, AKO2 and BKO1, BKO2.

Probing for RALA and RALB in selected SW403 clones resulted in complete absence of the targeted RAL isoform for all clones (Figure 4.6B). PCR clones AKO1, AKO4 and BKO3 and BKO4 were chosen to go forward, excluding the other clones that had extra PCR products present during the screening. These are subsequently referred to as AKO1, AKO2 and BKO1, BKO2 respectively.

Only two clones for each RAL isoform that had been confirmed by PCR screening (data not shown) were probed for RALA and RALB protein expression in HCT116 and SW48 cells (Figure 4.7). A clear absence of protein expression of the targeted RAL can be seen in all selected clones, whilst the other isoform remains unedited. All SW48 knockout clones come from a single parental cell line (Parental 3). The final knockout cell lines selected for study, along with their parental origin are outlined in Table 4.1.

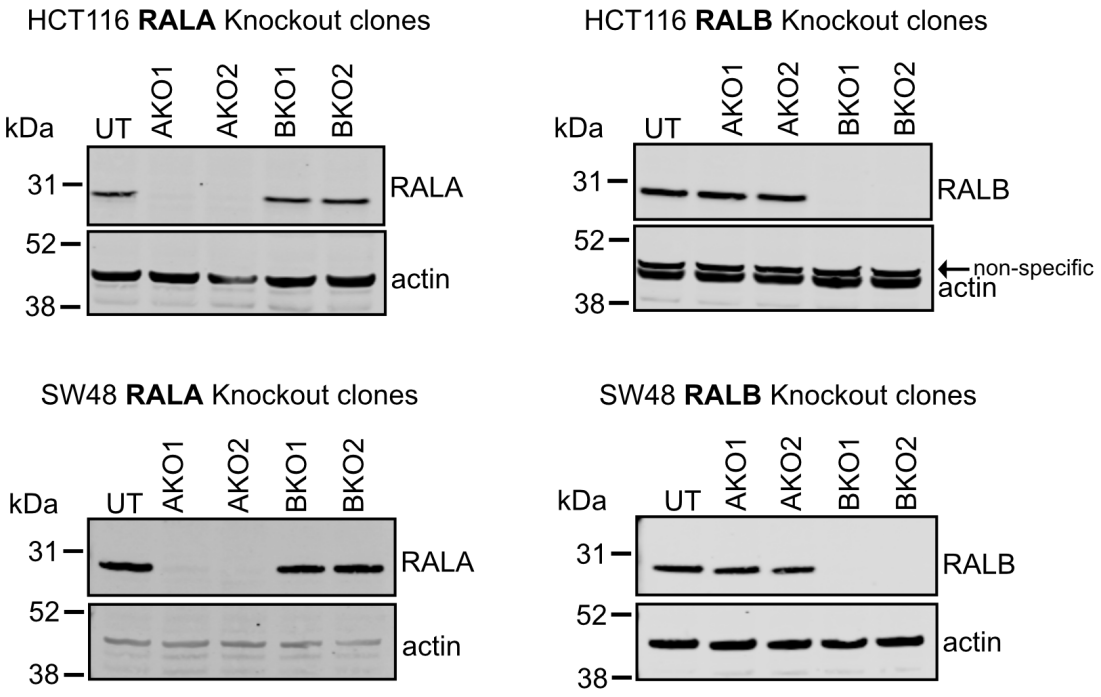


Figure 4.7 HCT116 and SW48 RALA/B knockout cell lines.
Isoform specific protein expression was absent in all positive clones identified by PCR, with protein expression of the non-target isoform unaffected in both HCT116 and SW48 cell lines. n=1

Table 4.1 Knockout cell lines and their parental origin.

KO Cell Line	RAL Target	Parental Origin	Updated Name
HCT116 AKO	RALA	PAR2	HCT116 AKO1
	RALA	PAR3	HCT116 AKO2
HCT116 BKO	RALB	PAR1	HCT116 BKO1
	RALB	PAR3	HCT116 BKO2
SW48 AKO	RALA	PAR3	SW48 AKO1
	RALA	PAR3	SW48 AKO2
SW48 BKO	RALB	PAR3	SW48 BKO1
	RALB	PAR3	SW48 BKO2
SW403 AKO	RALA	PAR1	SW403 AKO1
	RALA	PAR3	SW403 AKO2
SW403 BKO	RALB	PAR2	SW403 BKO1
	RALB	PAR3	SW403 BKO2
SW620 AKO	RALA	PAR2	SW620 AKO1
	RALA	PAR3	SW620 AKO2
SW620 BKO	RALB	PAR1	SW620 BKO1
	RALB	PAR3	SW620 BKO2
Final knockout clones selected for future experiments were given the updated name of KO1 and KO2. For three of the cell lines, all clones originated from 1 of 3 parental (PAR) clones with the intention of the two knockout clones for each RAL isoform descending from two different parental clones. SW48 KO clones were the only ones all with the same parental origin.			

4.3 RALA/B^{G23V} knock-in cell lines.

The homology directed repair (HDR) pathway is used to create specific point mutations or modifications into the genome. This requires the presence of a homologous DNA sequence that acts as a template for DNA repair. For both RAL isoforms, the point mutation of a guanine to valine at amino acid position 23 results in a constitutively active protein.

Although polymer and lipid-based transfection reagents have been suitable for initial validation of all gRNAs and generation of isoform specific knockouts, the efficiency of these reagents is very low in my panel of colorectal cell lines (section 3.2.1). Given that the efficiency of CRISPR/Cas9 mediated HDR is less than 1% itself, a lentiviral system was selected as a more efficient method of delivery of the CRISPR components (Figure 4.8A).

For the knockout cell strategy, the gRNAs and Cas9 endonuclease are contained within a single LentiCRISPR plasmid. Although the HDR DNA template can also be cloned into the same LentiCRISPR plasmid, all these components together would cause issues with packaging of the virus. For successful transduction to take place, the gRNA and HDR template need to be located within one plasmid without the Cas9. Instead, the Cas9 is integrated into the target cell genome and expressed separately within the cell.

Here, the gRNA and HDR template were subcloned into a pLeGO-EGFP/2A-Puro construct and inducible Cas9 cell lines were generated. The pLeGO was obtained from Addgene and modified by Dr Nicholas Harper, removing the Eu enhancer and replacing it with two multiple cloning sites to allow the gRNA and HDR template to be inserted into the same plasmid. The pLeGO construct also contains GFP and puromycin sequences for positive selection of transfected cells (Figure 4.9A).

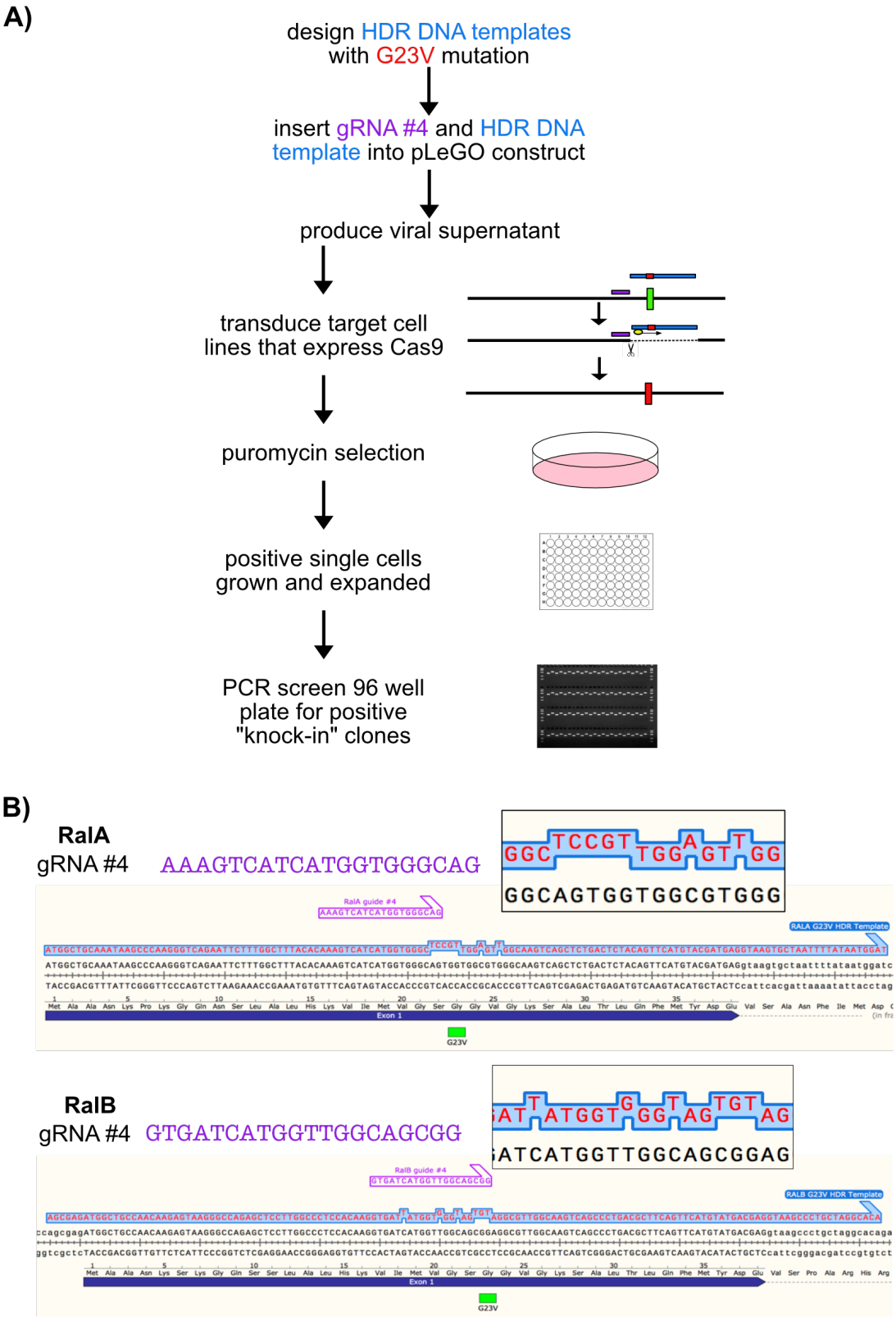


Figure 4.8 Generation of constitutively active G23V isoform specific RAL cell lines. **(A)** Strategy used to create isoform specific knock-in cell lines using the CRISPR/Cas9 system and lentiviral transduction. **(B)** HDR DNA templates for each RAL isoform including G23V mutation and silent mutations (raised letters) to aid screening process.

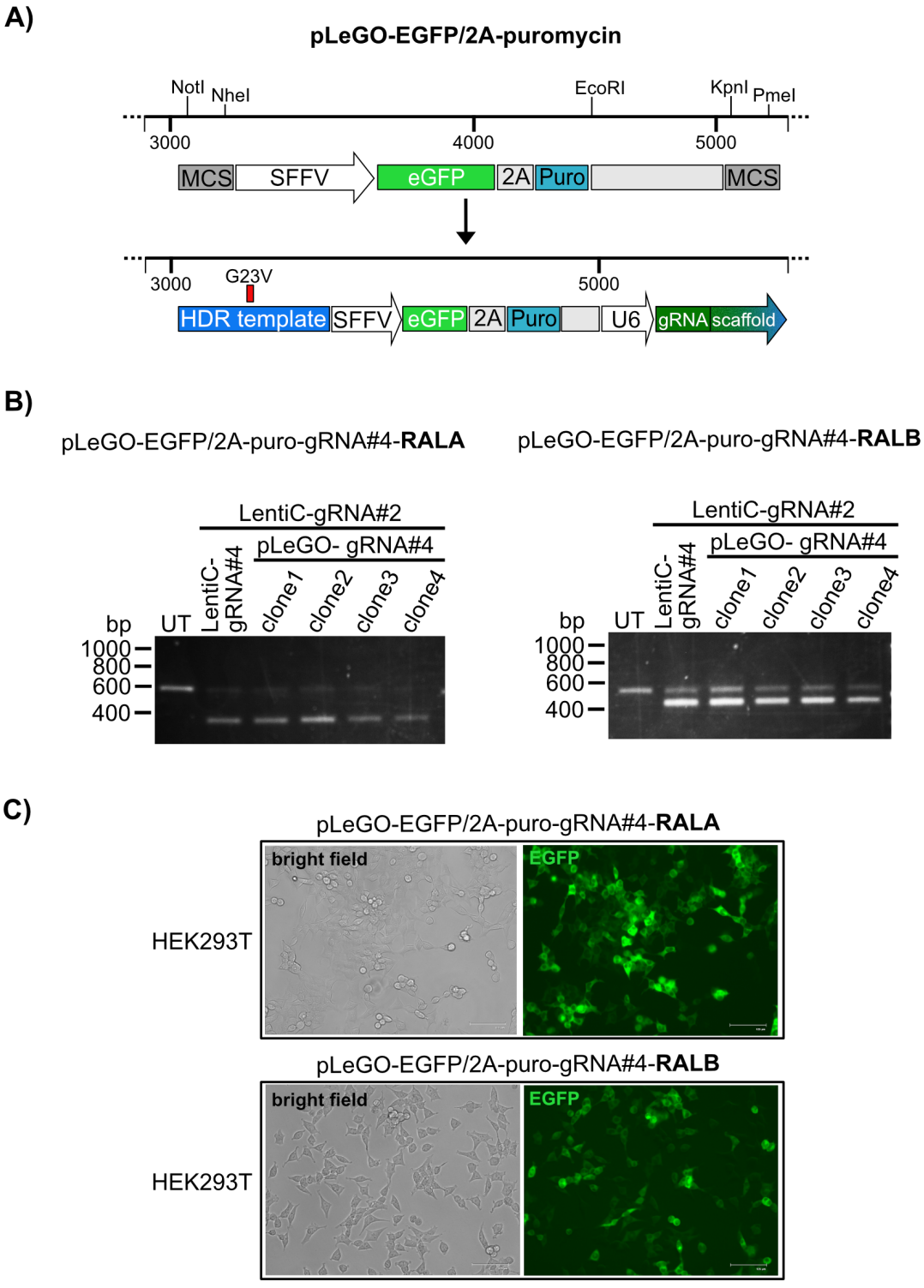


Figure 4.9 pLeGO-EGFP/2A-Puro construct for G23V knock-in cell lines.
A) Schematic of pLeGO construct containing both an enhanced GFP and puromycin selective markers. The G23V mutation HDR DNA templates for RALA and RALB were cloned in between *Not*I and *Nhe*I in the LH MCS, whilst the previously validated isoform specific gRNA#4 and a U6 promoter was cloned into the RH MCS. **B)** The efficiency of each gRNA#4 in the new pLeGO construct was validated by co-transfection with the corresponding LentiCRISPR-gRNA#2 (to provide Cas9) and screened by PCR for Cas9 cleavage and lower band shift. **C)** Transfection of pLeGO-EGFP/2A-Puro with gRNA#4 into HEK293T cells to confirm GFP expression to be used as a positive selection marker. GFP = green fluorescent protein, HDR = homology directed repair, MCS = multiple cloning site, SFFV = spleen focus forming virus, 2A = 18-22 amino acid-long peptides.

4.3.1 HDR DNA template design and insertion into pLeGO-EGFP/2A-Puro

For repair of the double strand break caused by Cas9 cleavage, a single stranded oligodeoxynucleotide (ssODN) containing the desired mutations was synthesised for each RAL isoform (Figure 4.8B). Both templates were designed to target the G23 amino acid and included changing bases of the RALA and RALB wild type sequence from a glycine to a valine. A number of silent mutations (Figure 4.8B) were also introduced into the HDR template that would aid in screening for positive knock-in cells via PCR as well as disrupting Cas9 PAM sequence binding and therefore prevent further gene editing following successful HDR.

For the HDR template to be cloned into the pLeGO construct, it needs to be converted into a double stranded oligodeoxynucleotide (dsOND) and requires longer homology arms than is needed for ssOND delivery. To generate this larger HDR template, two overlapping regions of genomic DNA surrounding the G23 codon were amplified by PCR using the ssOND HDR template as a primer to introduce the required mutations into the final template sequence (section 2.3.7).

4.3.2 Single gRNA#4 pLeGO plasmids modify RALA/B in HEK293T cells.

To generate G23V knock-in cell lines, a gRNA that is close to the desired mutation is preferred. For both RALA and RALB, gRNA#4 used in creating knockout cell lines, is close to the G23 position and has already been validated. Therefore, each gRNA#4 was excised from the LentiCRISPR constructs and subcloned into its own pLeGO vector (section 2.3.6).

Since the Cas9 gene is absent from the pLeGO constructs, 4 separate pLeGO-gRNA#4 plasmids for each RAL isoform were co-transfected into HEK293T cells with the complimenting single LentiCRISPR-gRNA#2 to provide gRNA#2 and Cas9 expression (Figure 4.9B). This allows the knockout screening method to be used to validate each gRNA#4 in the pLeGO construct.

A lower band shift (smaller PCR product) was observed in all co-transfections involving the LentiCRISPR-gRNA#2 and pLeGO plasmids containing either the RALA or RALB gRNA#4 (Figure 4.9B) compared to the untransfected. Co-

transfection of separate LentiCRISPR-gRNA#2 and LentiCRISPR-gRNA#4 was used as a control for successful Cas9 cleavage and deletion formation (Figure 4.9B).

Finally, successful transfection of the pLeGO plasmids with relevant gRNAs and HDR template can be seen through EGFP expression (Figure 4.9C).

4.3.3 Generation of doxycycline inducible Cas9 expressing SW403 and SW620 cells

To generate Cas9 expressing cells, my target cell lines were transduced with the integrating Cas9 expression plasmid pCW-Cas9-Blast (section 2.3.5). This plasmid contains a Cas9 that is under a reverse tetracycline controlled transactivator, along with a blasticidin selectable marker (Figure 4.10A). Following stable integration into the genome, expression of Cas9 can be induced with doxycycline drug treatment (Figure 4.10B).

Unfortunately, during antibiotic selection with blasticidin, in conjunction with single cell isolation, no viable HCT116 and SW48 clones were isolated. Only SW403 and SW620 cells survived treatment and were isolated and expanded. Following 48-hour treatment with doxycycline, several clones from each cell line were positive for Cas9 expression. Three clones in total from each cell line were selected for future knock-in cell line editing (Figure 4.10C).

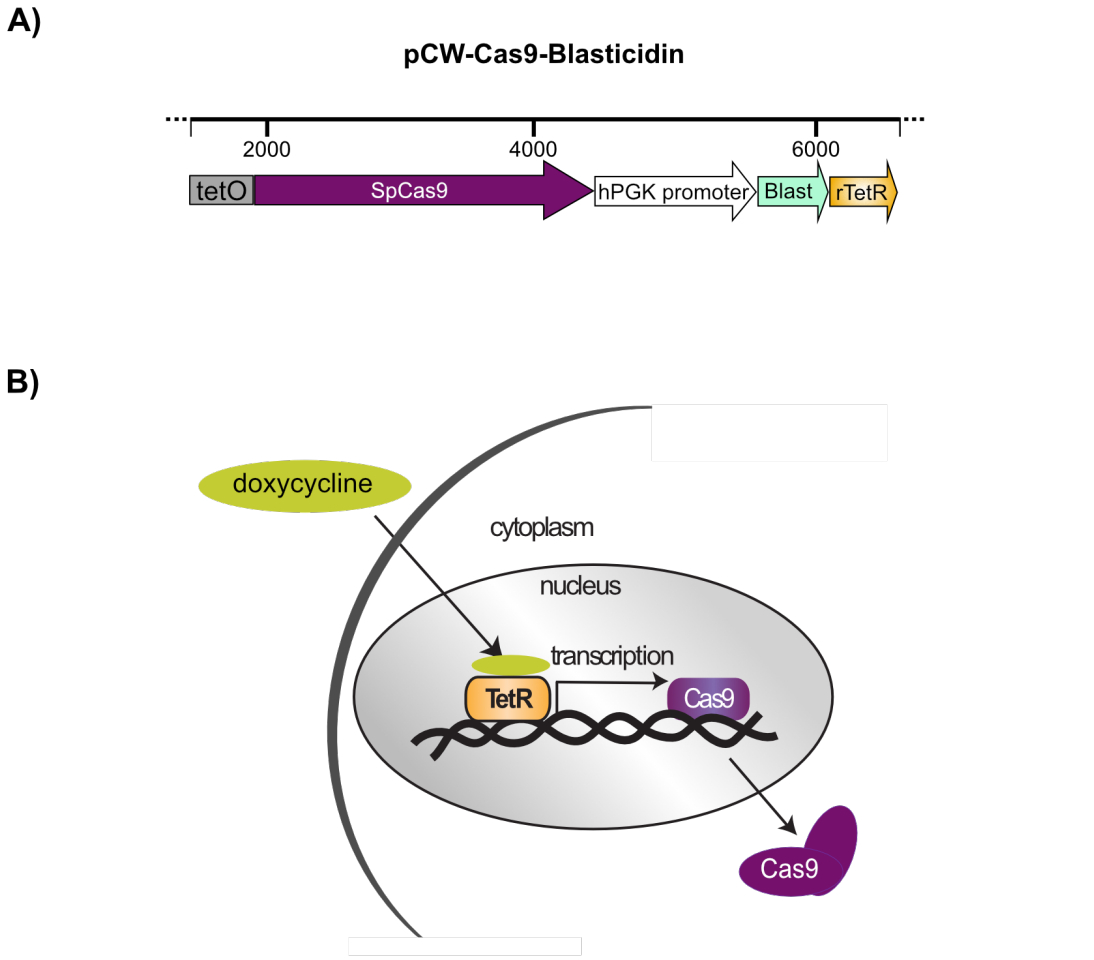


Figure 4.10 Generation of Cas9 inducible SW620 and SW403 cell lines.
A) Schematic of pCW-Cas9-Blast construct containing SpCas9 under a reverse tetracycline controlled transactivator with a blasticidin selectable marker. **B)** Expression of doxycycline inducible Cas9 following stable integration of the reverse tetracycline controlled transactivator (rtTA) dependent system. **C)** Expression of Cas9 can be seen in all transduced SW620 and SW403 clones following 48-hour treatment with doxycycline. Blast = blasticin, hPGK = human phosphoglycerate kinase, rTetR = reverse Tet repressor, SpCas9 = *Streptococcus pyogenes* Cas9, TetO = tetracycline operator.

4.3.4 PCR screening for identification of HDR.

Unfortunately, due to time constraints, RALA/B^{G23V} knock-in cell lines were not generated. If time permitted, the following protocol would be carried out to screen for successful HDR.

Once the SW403 and SW620 Cas9 inducible cells have been treated with doxycycline to induce enzyme expression, as well as transduced with either the pLeGO-EGFP/2A-Puro-gRNA#4-RALA or pLeGO-EGFP/2A-Puro-gRNA#4-RALB, positive cells will be selected through puromycin treatment and isolated as single cells into 96-well plates. Once the single clones have expanded, genomic DNA will be extracted from each well and will be screened by PCR to identify knock-in cells.

To screen for a successful knock-in of the desired HDR template, isolated DNA will undergo genotyping PCR using primers situated in Exon 1 in combination with primers designed to bind to the mutated sequence only. All primers were checked for specificity using NCBI Blast with the primer binding locations are shown in (Figure 4.11). The resulting PCR products for RALA using Primer Pair 1 (Gen 1F + Gen1R) and Primer Pair 2 (Gen2F + Gen2R) is 428 bp and 216 bp respectively. The resulting PCR products for RALB using Primer Pair 1 (Gen 1F + Gen1R) and Primer Pair 2 (Gen2F + Gen2R) is 292 bp and 294 bp respectively. Once knock-in cells have been identified they will be sent off for sequence verification.

A) **Genotyping Primers RALA**

Primer Pair 1
Gen 1F 5'-GCT GAT ACC CCA TTA TTG TCA GC-3'
Gen 1R 5'-GAC TTG CCA ACT CCA ACG G-3'

Primer Pair 2
Gen 2F 5'-GTG GGC TCC GTT GGA GTT G-3'
Gen 2R 5'-AAT CAA GTT ATG TAC AGA TGG CTC-3'

Genotyping Primers RALB

Primer Pair 1
Gen 1F 5'-GTG ACT TCC AGA GTT GTT GG-3'
Gen 1R 5'-ACG CCT ACA CTA CCC ACC-3'

Primer Pair 2
Gen 2F 5'-ATG GTG GGT AGT GTA GGC-3'
Gen 2R 5'-CGA TGC AA GCA GAG ATC TCT C-3'

B)

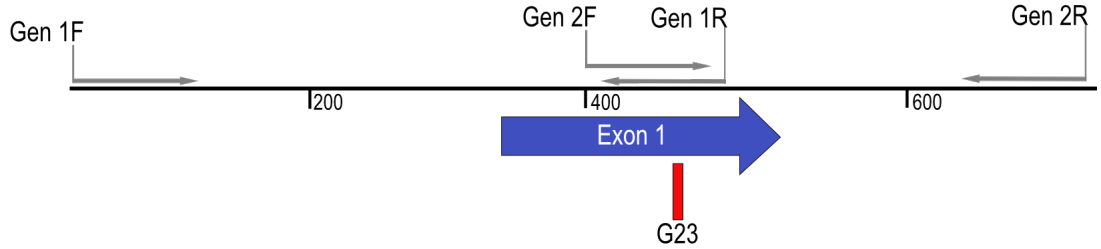


Figure 4.11 Genotyping primers for identification of knock-in mutations.
A) Genotyping primers were designed to bind to the mutated sequence (mutated bases underlined) to enable amplification of the inserted HDR template only, allowing for identification of knock-in clones. **B)** Position of genotyping primers required for identification of knock-in mutations. One primer is situated outside of the mutated region and the other designed to bind to the mutated knock-in sequence situated in exon 1 for both RALA and RALB. The resulting PCR products for RALA using Primer Pair 1 (Gen 1F + Gen1R) and Primer Pair 2 (Gen2F + Gen2R) is 428 bp and 216 bp respectively. The resulting PCR products for RALB using Primer Pair 1 (Gen 1F + Gen1R) and Primer Pair 2 (Gen2F + Gen2R) is 292 bp and 294 bp respectively.

4.4 Discussion

The aim of this chapter was to initially establish individual RALA and RALB knockout cell lines in the colorectal cells selected in chapter 3. This was done by validation of a CRISPR/Cas9-based strategy in HEK293T cells. Using multiple gRNAs for knockout generation resulted in larger DNA deletions compared to using a single gRNA. This has the advantage that larger sequence deletions can be easily detected through standard PCR instead of the more time consuming T7 endonuclease I assay, that recognises mismatched dsDNA. Due to the high transfection rate of HEK293T cells, validation of each gRNA was carried out before transfection into the target cell lines, as well as allowing subsequent identification of the most efficient gRNAs to generate separate RALA/B^{G23V} knock-in cell lines.

Four gRNA sequences per isoform were successfully cloned into separate modified LentiCRISPR plasmids, containing a functional Cas9 nuclease. Transfection of these constructs showed successful DNA cleavage by Cas9 in HEK293T cells (Figure 4.3). All Cas9 orthologues available for CRISPR gene editing are often limited by the placement of the PAM site required for gene targeting. This can be a problem for generating knockout cell lines but even more so for knock-ins, as gRNAs need to be designed adjacent to a PAM sequence. The SpCas9 used here relies on the more frequently appearing 5'NGG PAM site, increasing the likelihood of one being present in close proximity to our specific target sequence. Indeed, all gRNAs designed (4 for RALA, 4 for RALB), were able to target Exon 1 of their specific RAL isoform and resulted in a double stranded break which was detected by genotyping PCR (Figure 4.3). It has been shown that CRISPR/Cas9 editing efficiency decreases as the mutation to DNA cut distance increases (Paquet *et al.*, 2016). This means gRNA#4 for both RAL isoforms are expected to be the most efficient for knock-ins but gRNA efficiency can be tested by selecting positively transfected cells and performing quantitative PCR.

The different methods used to deliver the CRISPR/Cas9 system components into a cell population have their own advantages and disadvantages. The main strategies involve delivery of the Cas9 mRNA and gRNA as separate RNA

components, a plasmid-based system and finally, the delivery of the gRNA and Cas9 protein (Liu *et al.*, 2017). Here, only the plasmid-based system was used for both knockout and knock-in cell line generation.

The delivery of *in vitro* transcribed gRNAs, rather than through plasmid expression, reduces the risk of unwanted indels due to the short life of small RNA sequences. This method can be improved further by combining the gRNA delivery with either a Cas9 mRNA or Cas9 protein to bypass the use of DNA completely (Liu *et al.*, 2017). The synthesis of "ready-to-go" gRNAs has the advantage of avoiding time consuming cloning steps and cellular *in vitro* transcription. However, plasmid free delivery does not benefit from the addition of selection markers and is therefore not recommended for difficult to transfect cell lines such as these.

Using plasmids, like the lentiviral vectors used here, is limited due to restrictions in virus packaging capacity. Cas9 variants such as the saCas9 and St1Cas9 are smaller than the more popular SpCas9 and can help overcome this problem (Murovec, Pirc and Yang, 2017). However, with the HDR DNA template required for knock-in cell lines in addition to the 4.2 kb SpCas9 endonuclease, delivery into my panel of cell lines presents a major challenge when using the CRISPR/Cas9 system. Consequently, using a lentiviral vector for the gRNA and HDR template along with separate Cas9 expressing cells ensures a higher transduction efficiency of the colorectal cell lines.

A disadvantage to the stable integration and expression of Cas9 in target cells is the occurrence of off-target gene editing. It has been shown that Cas9 can induce mutations that differ from the target site (Chew *et al.*, 2016; Fu *et al.*, 2013) and that an increase in off-site Cas9 cleavage correlates with high Cas9 expression levels (Fu *et al.*, 2013). To avoid continuous Cas9 expression, a plasmid containing an inducible promoter was used to control expression and keep levels down.

The tetracycline regulation system used here, is very sensitive and commonly used in mammalian cell culture, mice as well as other species (Howe *et al.*, 1995). Whilst inducible gene expression plasmids allow temporal and quantitative control over gene expression, they can be limited by a number of factors such as the toxicity of the inducing agent and low induction expression depending on the expression system used. However, the doxycycline

concentration of 8 µg/mL used here was sufficient at inducing Cas9 expression with no toxic or detrimental effects seen on cell health. Alternatively, a self-limiting SpCas9 that removes the nuclease from edited cells, preventing the incidence of further off-target modifications could also be used (Petris *et al.*, 2017).

For introducing small mutations such as the ones required here, an HDR DNA template in the form of a ssODN is preferable to larger donor plasmids (Ran *et al.*, 2013b). The HDR DNA templates used to introduce the G23V point mutation in both RAL isoforms were originally designed as a 138 bp and 139 bp ssODN respectively. For small mutations, a ssODN should be around 100-150 bp total homology with the mutation introduced in the middle and flanking sequences of at least 40 bp on each side (Chen *et al.*, 2011a; Ran *et al.*, 2013b). Silent mutations were included around the G23 target site to allow for positive identification by PCR using mutation specific primers. The addition of silent mutations also destroys the PAM sequence, preventing any further Cas9 cleavage following the occurrence HDR mediated editing.

Homology arm length, symmetry and chemical modifications of the HDR DNA template can affect the efficiency of the HDR pathway. Using an unmodified template with asymmetrical homology arms has been shown to increase HDR efficiency (Richardson *et al.*, 2016). However, the use of a phosphorothioate modified template with symmetric arms has been shown to improve knock-in efficiency more so than an unmodified template (Renaud *et al.*, 2016) and may be worth consideration.

Whilst these optimisations are suitable for transfection involving a Cas9/gRNA-expressing plasmid and separate ssODN DNA template, as previously mentioned, the transfection efficiency of these colorectal cell lines is poor and therefore delivery of components through viral transduction is preferred. With that in mind, delivery of the gRNA, as well as the HDR template required a separate pLeGO plasmid that contained an additional selection marker. A selection marker is not necessary when an introduced mutation changes the cell phenotype, as the edited subpopulation can easily be isolated via flow cytometry. Since this is not known to be the case for the introduction of the RALA/B^{G23V} mutation, a GFP selection marker was used for easy monitoring of transduction efficiency and positive cell selection.

An integrase defective lentiviral vector (IDLV) was used to package the gRNA and HDR template. Most IDLVs have mutations introduced in the catalytic triad of the integrase region, such as a residue D64 mutation seen here, that results in the inactivation of the integrase enzyme. The use of IDLVs compared to standard lentiviral vectors has been shown to have transient expression with a decreased capacity to induce off-target mutations (Ortinski *et al.*, 2017). In this instance it was also used to prevent integration of the GFP sequence into the cell genome preventing the future use of fluorescent analysis.

If the transduction efficiency of the integrase deficient virus is very low, the GFP selection marker can be used to sort cells by fluorescence-activated cell sorting (FACS). However, whilst this will result in an enrichment of transduced cells, along with puromycin selection, it may increase the stress on the cells and affect overall viability. Furthermore, it increases post-selection recovery time which may lead to contamination and loss of cells. Therefore, for future experiments, if the viral transduction efficiency is too low, the IDLV will be exchanged for a standard packaging vector such as psPAX2 and the GFP selection marker removed from the pLeGO plasmid to prevent integration of the fluorescent marker into the host genome.

Knock-in efficiency via the HDR pathway is reported as being as low as 1% (Wang *et al.*, 2015) meaning that only a small percentage of transduced cells will contain the RALA/B^{G23V} mutation. HDR can be enhanced by suppressing key proteins in the NHEJ repair pathway. Inhibition of Ku70, Ku80 and DNA ligase IV can increase HDR efficiency by up to 4-5-fold (Chu *et al.*, 2015). Exogenous expression of the homologous recombination protein Rad52 can also increase HDR and when combined with DNA ligase IV inhibition, using the small molecule inhibitor Scr7, it can improve repair efficiency up to 40% (Shao *et al.*, 2017). However, these strategies have been tested using small inserts on few gene loci, so it is not clear whether they would enhance a RALA/B^{G23V} knock-in. Furthermore, inhibition of the NHEJ pathway alone may cause toxicity. Another strategy to increase HDR efficiency would be to use double nicking. This method involves a Cas9 D10A nickase that creates a single-strand rather than double-strand DNA break. When used with two gRNAs that target opposite DNA strands, each gRNA pair has been shown to successfully induce HDR at frequencies higher than those of single gRNA with

WT Cas9 (Ran *et al.*, 2013a). This method has also been shown to reduce off-target activity.

Finally, once RALA/B^{G23V} knock-in clones have been generated, the whole RAL isoform gene will be sequenced to determine if any additional mutations surrounding Exon1 have been inserted. Any additional mutations could result in changes in protein conformation and activity affecting functional analysis. Moreover, any other suspected off-target modifications would require whole genome sequencing.

Chapter 5: Phenotypic profiling of cell lines

5.1 Introduction

Generation of a panel of isoform specific knockouts now allows us to look at RAL isoform signalling. It has been demonstrated that signalling responses vary between the different RAL isoforms in both normal biological functions as well as tumour cell progression. RALB but not RALA has been shown to act as a regulatory switch, promoting biogenesis of autophagosomes through direct engagement of the Exo84 effector, following cellular starvation in HEK293T cells (White et al. 2011). Knockdown of RALA and RALB leads to opposing effects on the development of tight junction barrier function in MDCK II dog cells (Hazelett et al. 2011). Constitutively active RALA but not RALB promoted basolateral delivery of secretory vesicles (Shipitsin & Feig. 2004). Knockdown of RALA but not RALB affects mitochondrial fission and morphology in human HEK-TtH cells (Kashatus et al. 2011). Different RAL requirements have also been seen in embryo development, where RALB null mice are viable but RALA null leads to exencephaly and embryonic lethality (Peschard *et al.*, 2012).

The idea that these two highly related proteins could also serve opposing functions in oncogenic RAS processes has been previously described. Knockdown of RALA expression impeded and, in some cases, prevented tumourigenesis of human cancer cells (Lim et al. 2005). Interfering RNA stable suppression of RALA has also been shown to reduce anchorage-independent growth in colorectal cancer cell lines, whilst suppression of RALB had a contrasting effect of enhancing soft agar colony size and frequency (Martin et al. 2012). Despite these opposing activities, both RAL isoforms required interaction with the effector RALBP1 (RLIP76), as well as members of the exocyst complex. However, it was determined that RALA required exocyst component Sec5 but not Exo84 and vice versa for RALB, highlighting the differences in the direct protein-protein interactions involved in RAL signalling. Again, RALA but not RALB is required for anchorage-independent growth *in vitro* and tumourigenic growth *in vivo*, in human pancreatic carcinoma cells (Lim et al. 2006), whereas lung metastasis was more RALB dependent (Lim et al. 2006).

Any divergent functional roles are thought to be due, in part, to the differences in subcellular location of each isoform (Gentry *et al.*, 2015). Furthermore, even though RALA and RALB share a high degree of amino acid sequence identity (Gentry *et al.*, 2014), distinct biological roles have also been largely attributed to their differing C-terminal hypervariable regions (Shipitsin and Feig, 2004; Lim *et al.*, 2005). The varied posttranslational modifications of the C-terminal of RALA and RALB, also result in different RAL isoform regulation and therefore function (Gentry *et al.*, 2015).

Although these observations suggest a contrast in RAL isoform functionality, many in an appropriate cell model, these studies rely on temporary suppression or overexpression of RAL proteins. Additionally, whilst RALA induced anchorage independent growth was found to be independent of KRAS mutation status (Martin et al. 2012), somewhat surprisingly, the influence of RAS status and whether it presents as a cofounding factor in RAL signalling has been largely ignored. Potentially due to insufficient cell model availability and lack of knowledge of RAL signalling in general, studies have focused on understanding interactions downstream of RAL and not its regulation by RAS. Ultimately, whether the status of RAS is important in RAL function or whether RAL is even involved in oncogenic RAS signalling is still unclear.

The RAL knockout cells generated in chapter 4 allow us to address RAL isoform activity and function at an endogenous level in an appropriate model for RAL study. Unlike other RAL studies, profiling a larger panel of knockout cell lines will not only account for clonal differences but also any variations between cell lines. Furthermore, since RAS is expected to be engaged in RAL signalling, this panel of colorectal cell lines includes different KRAS mutation statuses, allowing for direct comparison and therefore collection of a more robust profile of RAL biology.

With the recent development of various omic tools, it is possible to systematically characterise RAL signalling networks without altering location or protein structure/activity through modifications. There are a variety of omic techniques that can be used to cover gene expression right through to protein analysis, each with their own benefits.

Various microarray techniques can be used to profile the genome. In fact, gene expression profiling via microarray following separate RALA and RALB

knockdown in the UMUC-3 bladder cancer cell line has already revealed RREB-1 as a novel transcriptional RAL effector (Oxford et al. 2007). However, RNA sequencing (RNAseq) is a more sensitive method in detecting gene expression changes especially when at very low levels. It is also more accurate at detecting extremely abundant genes and has a wider dynamic range than microarray (Zhao et al. 2014). RNA samples can be easily obtained and reproducibly measured in a high-throughput manner using a variety of sequencing technologies (Marioni et al. 2008 and Elliot et al. 2009). Indeed, the advent of next generation sequencing has revolutionised transcriptomics and established RNAseq as the preferred method for the study of gene expression (Finotello et al. 2014). The expression levels of RNA can be used as a proxy for protein abundance, but due to various regulation stages, the presence of specific RNA products does not necessarily guarantee a functional end product. Therefore, the use of additional omic tools looking at the proteome, to see if any gene expression changes generate meaningful protein changes, is desirable.

Whilst whole proteome investigation provides direct insight into actual protein level changes, analysis of large volume of data can be time consuming. Moreover, protein levels in themselves don't necessarily tell you if something is changing or address any post translational modifications. Signalling status can be determined by lower throughput methods such as western blot but a more useful approach is to focus on families of proteins using an assay such as Reverse Phase Protein Arrays (RPPA) (Spurrier, Ramalingam and Nishizuka, 2008; Baldelli *et al.*, 2017). RPPA is not only relevant for cancer pathway study, but it also provides the ability to infer protein network responses of hundreds of proteins/phospho-proteins and capture the state of key signalling transduction pathways.

By conducting unbiased large-scale screens on these new RAL knockout cell models, a larger data set of RAL signalling networks can be collected. This will provide a broad understanding of the changes that have occurred in these cell lines due to gene knockout and pin any changes, where possible, onto direct RAL dependent signalling.

5.1.1 Aims and Objectives

The aim of this chapter is to determine the effect of single RAL isoform deletion on the RAL pathway and wider cell signalling networks.

Objectives:

- Determine activity changes in the remaining RAL isoform using the active RAL pulldown assay used in chapter 3.
- Profile the transcriptome of parental and RAL knockout lines using RNA sequencing to determine changes in mRNA expression.
- Profile the proteome of the same cell lines using RPPA to see if changes extend to the protein level.
- Validate proteins of interest by western blotting.

5.2 Results

5.2.1 Isoform specific RAL KO activity assays

Following pulldown of endogenous active RAL in the unedited parental cell lines in chapter 3, I wanted to see what effect knockout of each RAL had on the activity of the remaining RAL isoform. The same commercial pulldown method for isolating active RAL was used as outlined in chapter 3. To profile RAL isoform activity in the KO vs PAR cell lines, I used 2 minutes of EGF stimulation since this corresponds to the average peak RAL activity (Figures 3.2-3.3). Probing for phospho-ERK was also carried out to confirm stimulation of RAS by EGF. RAL activity assays were repeated 3 times (individual repeats in supplementary material) and the RAL activity for each clone was quantified.

The results for individual cell lines are shown in Figures 5.1-5.4 and a summary is shown in Figure 5.5. Activity of the remaining RAL isoform was recorded in all edited clones, across all cell lines, and no visible activity of the "knocked-out" RAL could be seen (Figure 5.1A-5.4A). Whilst activity varied between RAL KO clones within cell lines, activity of the remaining isoform always trended in the same increased direction. For RALA, a small increase was observed in some of the BKO samples, whereas RALB activity was significantly increased in the RALA KOs, across all cell lines except SW48. When looking at individual cell lines, the increase in RALA activity in the SW48 and SW620 RALB KO clones was not as pronounced as the HCT116 and

SW403 cell lines (Figure 5.2 and 5.4) but could still be seen clearly for clone BKO2 in the SW620 cell line. Overall, a clear compensatory increase in the other isoform was seen in all cases except the SW48 cell line (Figure 5.2A). No change in overall protein expression of either RALA or RALB was detected in any of the cell lines (Figure 5.1A-5.4A). Therefore, it is inferred that loss of one isoform affects the activity but not the protein expression levels of the other.

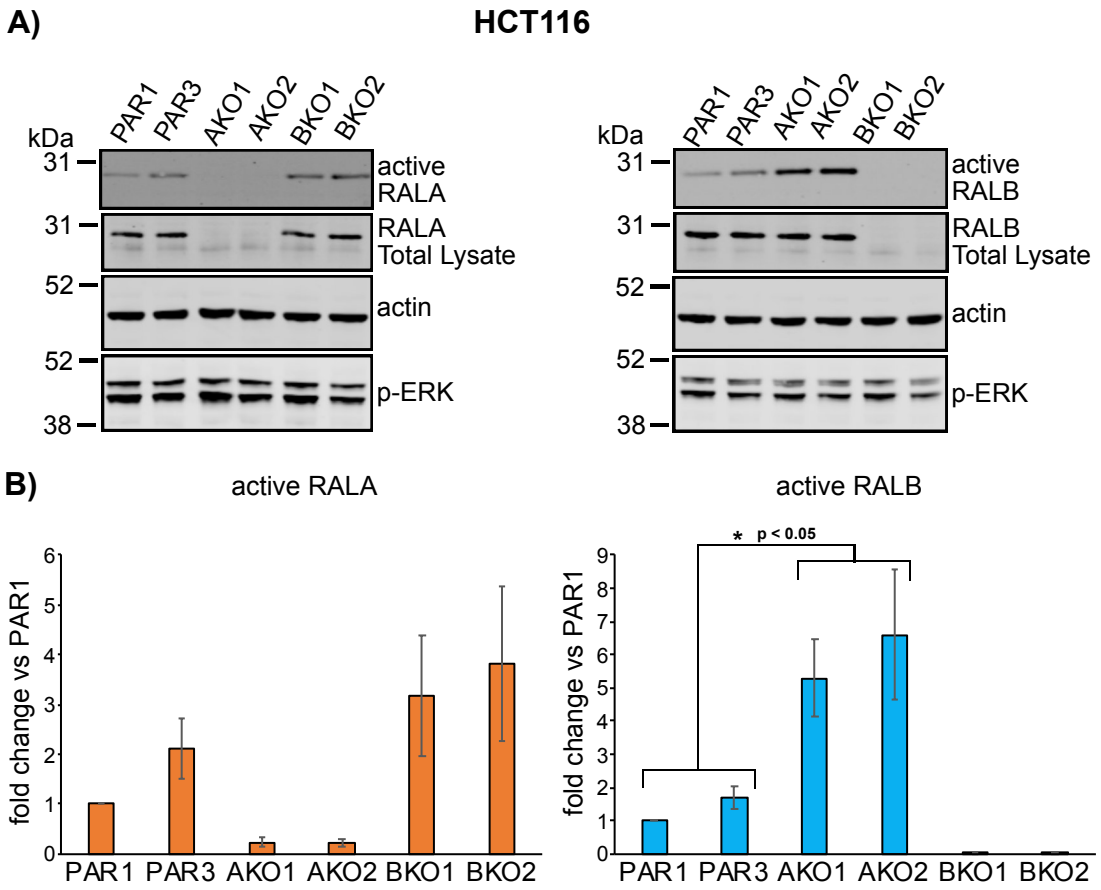


Figure 5.1 RAL activity in HCT116 isoform specific knockout cell lines.
A) Following treatment with 50 ng/mL EGF for 2 minute, HCT116 RALB KO clones, BKO1 and BKO2, displayed a small increase in RALA activity compared to the parental (PAR1 and PAR3). An obvious increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2). Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate. **B)** Quantification of RAL activity confirmed a significant increase ($p < 0.05$) in RALB activity vs PAR in the RALA KO clones. Data is presented as a fold change and represents an increase or decrease in activity compared to PAR1; fold change ± SE of n = 3 biological replicates.

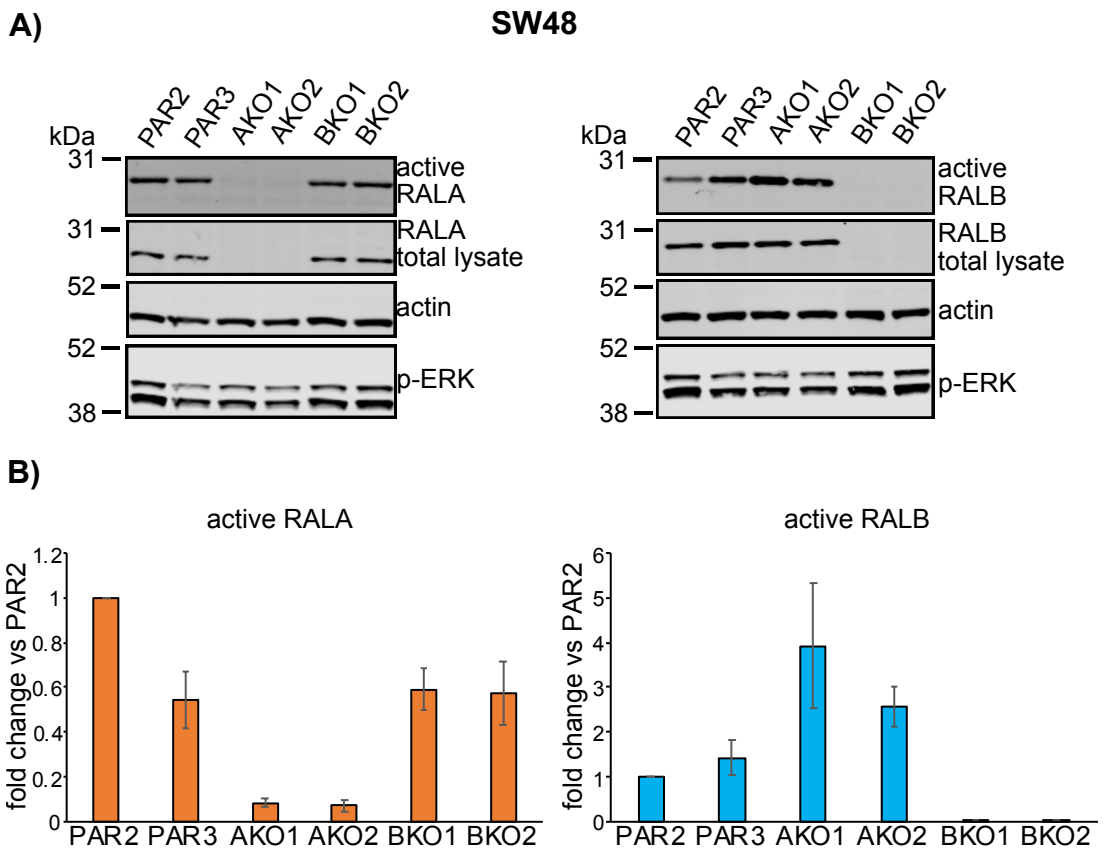


Figure 5.2 RAL activity in SW48 isoform specific knockout cell lines.
A) Following treatment with 50 ng/mL EGF for 2 minute, SW48 RALB KO clones, BKO1 and BKO2, displayed similar RALA activity compared to the parental (PAR2 and PAR3). An obvious increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2). Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate. **B)** Quantification of RAL activity confirmed no significant increase in activity in the KO clones. Data is presented as a fold change and represents an increase or decrease in activity compared to PAR2; fold change ± SE of n = 3 biological replicates.

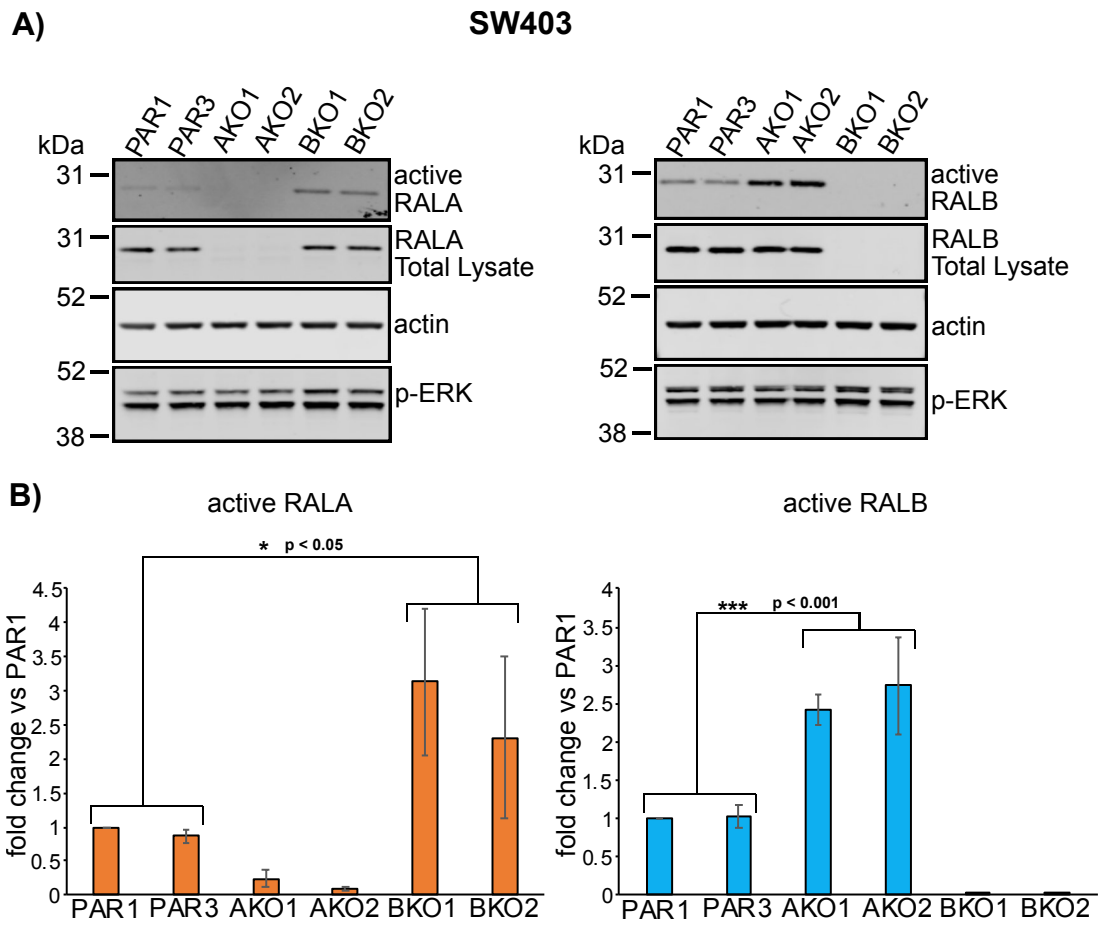


Figure 5.3 RAL activity in SW403 isoform specific knockout cell lines.
A) Following treatment with 50 ng/mL EGF for 2 minute, SW403 RALB KO clones, BKO1 and BKO2, displayed a small increase in RALA activity compared to the parental (PAR1 and PAR3). An obvious increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2). Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate. **B)** Quantification of RAL activity confirmed a significant increase in RALA activity vs PAR in the RALB KO clones ($p < 0.05$) and a significant increase in RALB activity vs PAR in the RALA KO clones ($p < 0.001$). Data is presented as a fold change and represents an increase or decrease in activity compared to PAR1; fold change ± SE of n = 3 biological replicates.

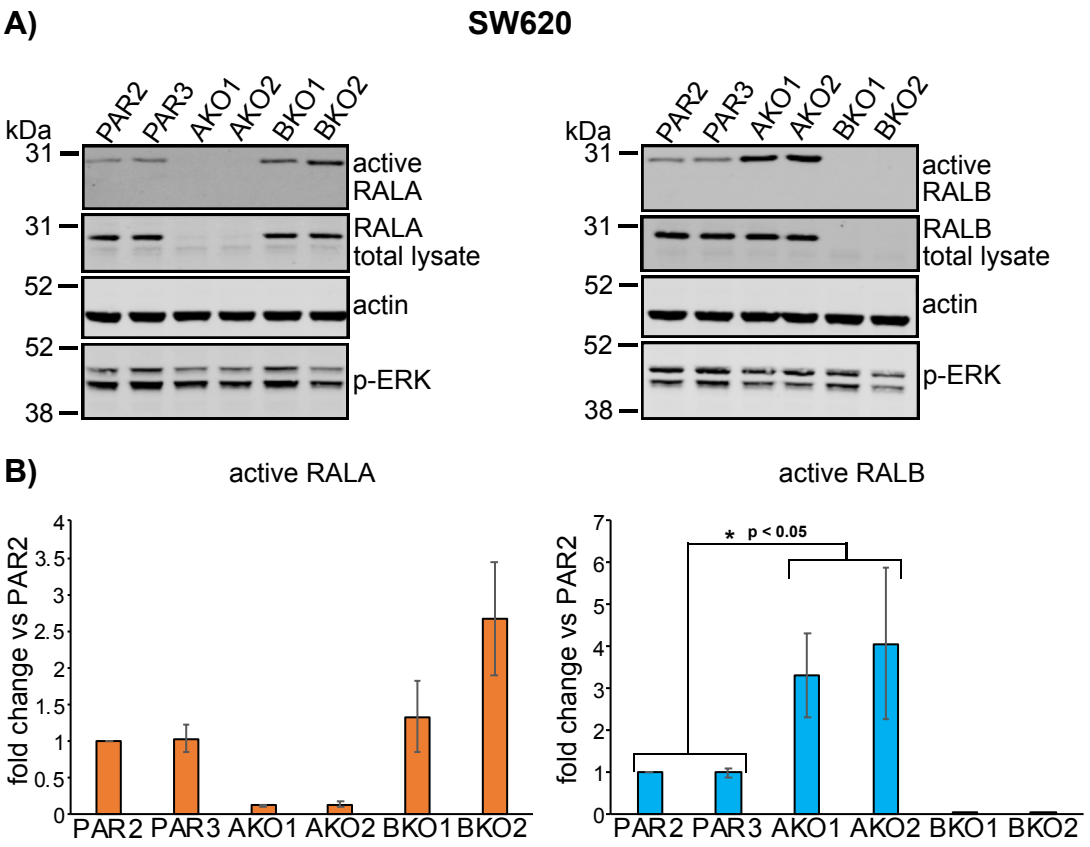


Figure 5.4 RAL activity in SW620 isoform specific knockout cell lines.
A) Following treatment with 50 ng/mL EGF for 2 minute, SW620 RALB KO clone BKO2, displayed a small increase in RALA activity compared to the parental (PAR2 and PAR3). A significant increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2). Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate. **B)** Quantification of RAL activity confirmed a significant increase in RALB activity vs PAR in the RALA KO clones ($p < 0.05$). Data is presented as a fold change and represents an increase or decrease in activity compared to PAR2; fold change ± SE of n = 3 biological replicates.

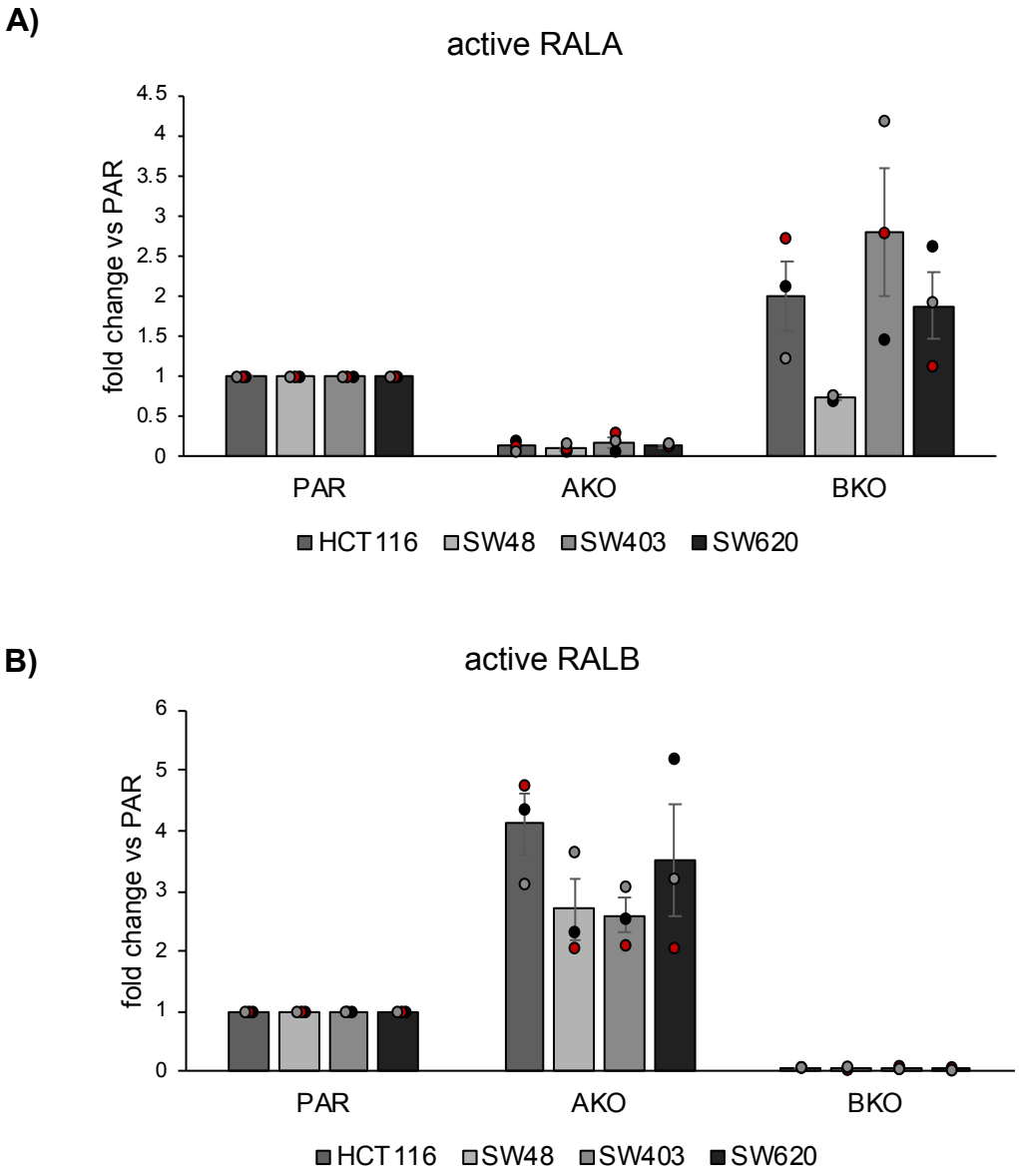


Figure 5.5 RAL activity in RAL knockout cell lines vs PAR across all cell types.
A) Quantification of RALA activity across all PAR and RAL KO cell types. The RALA activity varies across all cell types in the BKO clones and parental. SW403 display the highest RALA activity out of the cell types. **B)** Quantification of RALB activity across all parental and RAL KO cell types. Again, RALB activity is varied between cell types in the AKO clones with the highest activity observed in the HCT116 RALA KO clones. Data is presented as a fold change and represents an increase or decrease in activity compared to PAR; fold change ± SE of n = 3 biological replicates. Circles (black, grey and red) represent the individual fold change each repeat.

5.2.2 Transcriptome analysis of PAR and RAL Knockout cell lines using RNA sequencing

Since RAL isoform specific knockouts were first generated in the SW403 and SW620 colorectal cell lines, these were selected to undergo RNA sequencing. This included two parental (PAR), two RALA KO and two RALB KO clones for each cell type, giving a total of twelve samples.

5.2.2.1 Gene expression profiles

In order to reflect the correlation of gene expression between all twelve samples, the Pearson correlation coefficient of expression levels was calculated between each two samples. RNAseq gene expression profiles are different between cell types regardless of RAL status (Figure 5.6A).

All RNAseq samples then underwent hierarchical clustering by the expression level of all genes (Figure 5.6B). This directly reflects the relationship between each two samples, where the closer the samples are to each other, the more similar the expression level. For the RNAseq, the two cell lines remain as distinct branches separate from each other, whether they be parental or knockout. For SW403, RAL KO samples had a more similar expression profile with each other than compared with the parental, regardless of RAL isoform targeted. The parental cell lines can be seen grouped together in both cell types, with the RALA KO clones paired off in SW620, with RALB KO2 branching off the main knockout branch. Only RALB KO1 from the SW620 cell type was clustered with the parental clones. This could reflect the similar levels of RALA activity of clone BKO1 vs parental seen in Figure 5.4 of chapter 5.

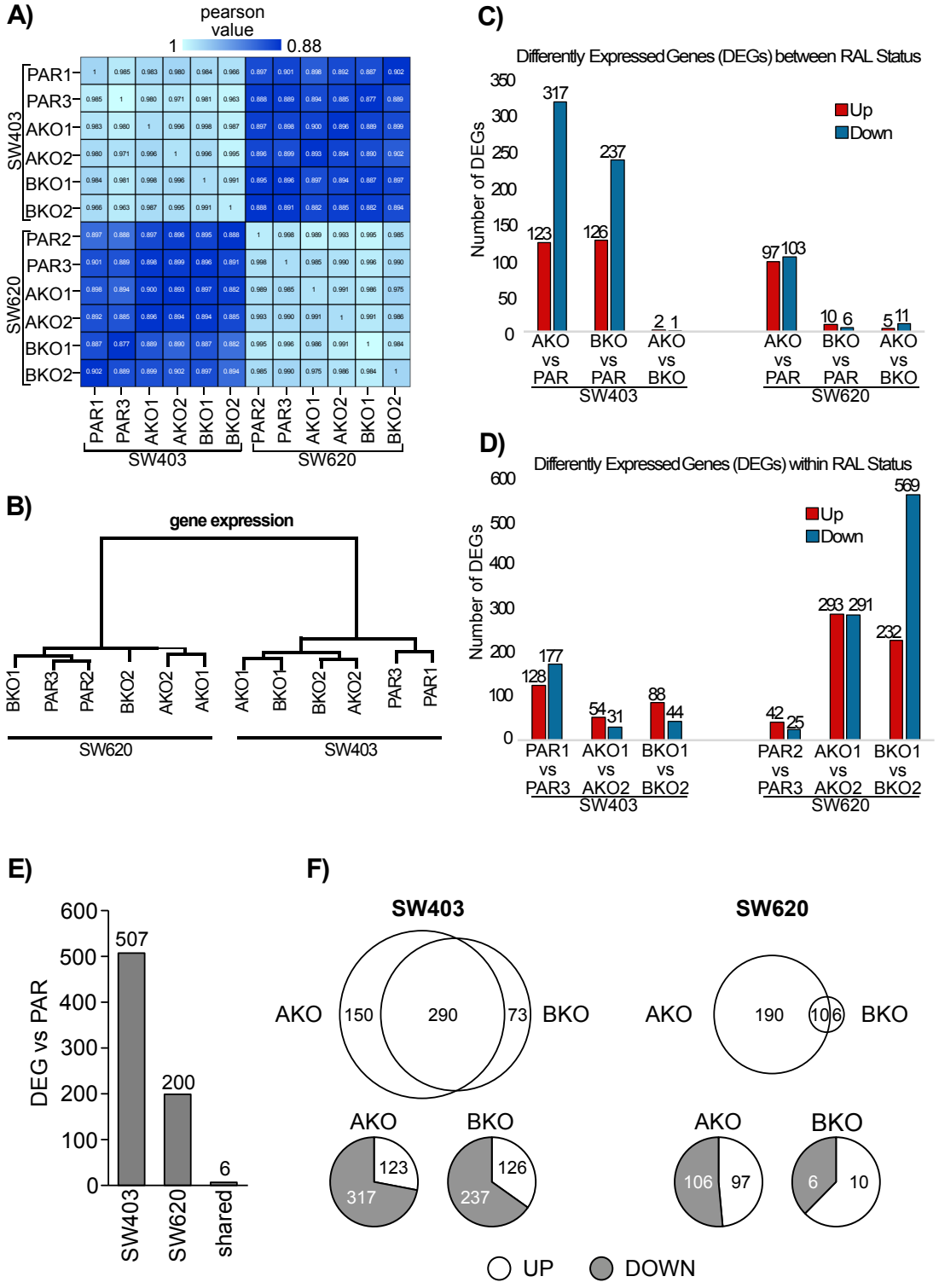


Figure 5.6 Correlation between samples and comparison of Differentially Expressed Genes (DEGs) from RNAseq data analysis.
A) Pearson correlation coefficient values for all gene expression levels between each two samples. **B)** Hierarchical clustering of the expression level of all genes between each sample. **C)** Number of DEGs between RAL status in SW403 and SW620 cell lines. **D)** Number of DEGs within two samples of the same cell type with the same RAL status for SW403 and SW620 cell lines. **E)** Total number of DEGs in single RAL knockout (RAL KO) cell lines compared to the parental (PAR). Only six genes are shared between the two cell types (SW403 and SW620) that are significantly different. **F)** Breakdown of DEGs into RALA KO, RALB KO and shared within each cell type. Breakdown of RALA KO and RALB KO DEGs according to whether they increase or decrease compared to PAR.

5.2.2.2 Identification of shared differentially expressed genes (DEGs)

Based on the gene expression level and the parameters previously mentioned (fold change ≥2.00 and adjusted P value ≤0.05), the differentially expressed genes (DEGs) between RALA/B KO and PAR samples can be identified.

Overall, more DEGs were identified in SW403s, with a total of 513 DEGs present in the RAL KOs vs PAR, than in the SW620 cell line, which had a total of 206 DEGs (Figure 5.6E). Although the number of total DEGs differs between the SW403 and SW620 clones (Figure 5.6E), more differentially expressed genes were detected in the RALA KOs for both cell types, suggesting it is the more dominant of the two RAL isoforms when it comes to RAL network signalling and potentially function. For most of the DEGs in the SW403 cell line, gene expression was decreased in both RAL KOs in comparison to the parental, whilst a more even distribution of increased and decreased DEGs was observed in the SW620 samples (Figure 5.6C and Figure 5.6F). This suggests that different cell types might have different preferences for RAL isoform signalling. The small number of differentially expressed genes between RAL isoform samples, implies the two isoforms are involved in similar gene interactions and same signalling networks. Very few if any genes were identified where they increased in one isoform KO and decreased in the other (Figure 5.7). Again, this highlights the idea that there is some functional compensation within the two RAL isoforms.

Although DEG identification between RAL status did not include genes that were only differentially expressed in one of the two knockout clones, it is still important to consider any gene expression differences within the two RAL KO samples. For the SW403s, PAR had the most DEGs present between the two clones (Figure 5.6D). In comparison, the knockout clones for SW620 have more DEGs present between isoform specific KO clones than the PAR clones (Figure 5.6D). This might explain the discrepancy between the higher total of DEGs identified in the SW403 cell type compared to the lower numbers recorded for SW620. Significant changes in gene expression between two samples with the same RAL status, emphasises the presence of clonal variation and how two samples can have different expression profiles and consequently differing cell signalling networks, regardless of gene editing. This is an important consideration when it comes to future data analysis and

interpreting RAL isoform specific function from clones with varied gene expression.

Only 12 DEGs are shared between SW403 and SW620 (Figure 5.7A), with different expression trends seen in genes PTHLH (parathyroid hormone like hormone), PLXNA2 (plexin A2), IGFBP2 (insulin like growth factor binding protein 2) and PLAU (plasminogen activator, urokinase). The enzyme CHST3 (carbohydrate sulfotransferase 3), is the only shared gene that is increased across all KO samples vs PAR and is the only shared gene that is classified as a DEG under a different RAL status (SW620 AKOs and SW403 BKOs), whilst trending in the other RALKO. Surprisingly, only 6 DEGs displayed decreased expression between the two cell lines (Figure 5.6E and 5.7A). PLK2 (polo-like kinase 2), MYZAP (myocardial zonula adherens protein), EREG (epiregulin), GJB3 (gap junction protein, beta-3), ETS1 (ETS proto-oncogene 1) and PRSS3 (serine protease 3) are all downregulated in the RAL KO samples (Figure 5.7A). This reflects the lack of gene expression correlation as a whole seen in Figure 5.5 and how different these two cell lines are at the genome level.

To determine if the top DEGs for one cell type also trended in the other cell type, heatmaps were generated comparing RALA specific, RALB specific and RALA and RALB specific DEGs between SW403 and SW620 samples (Figure 5.7B). It should be pointed out that although the DEGs from both cell lines were selected for comparison against each other, the higher total number of DEGs for the SW403 mean a lack of expression change is observed on the SW620 side. In fact, no DEGs were present in both AKO and BKO clones in the SW620 cell line. Overall, there was little similarity in the expression trends between the two cell lines. Many of the DEGs for one cell line, were either absent, showed no change or even showed an opposite trend in expression for the other type (Figure 5.7B). This lack of trending gene expression further highlights the difference in genomic profiles between cell lines at the RNA level.

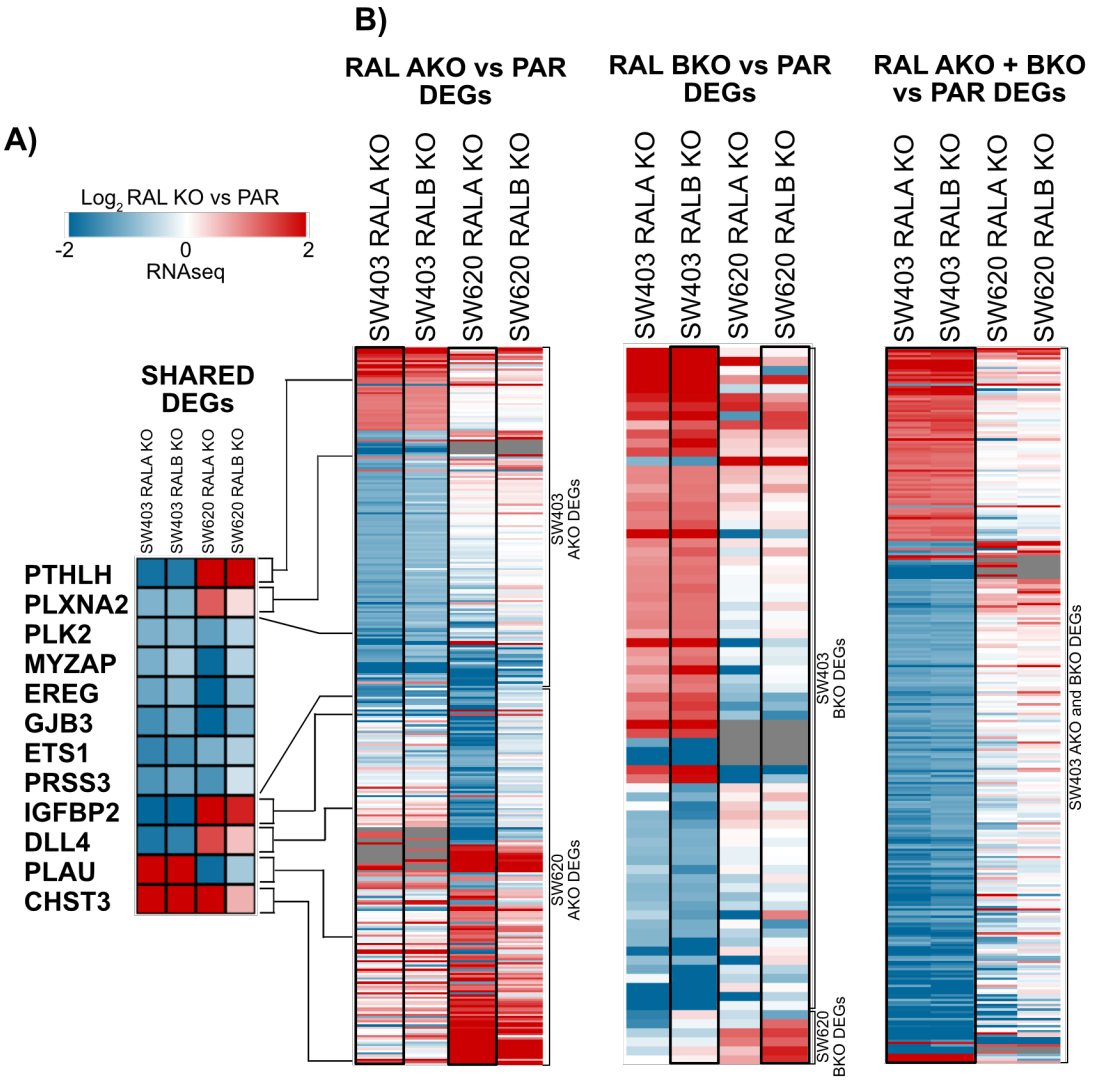


Figure 5.7 Differentially Expressed Gene (DEGs) trends.
Comparison of RALA specific, RALB specific and both RALA and RALB DEGs between SW403 and SW620 cell lines displayed little similarities between the two cell lines. DEGs in both the AKOs and the BKOs were only present in the SW403 cell line. The difference in gene expression profiles between the two cell lines is highlighted by the lack of expression change for many genes in one cell line but are significant for the other. Total lack of expression (grey) for some DEGs can also be seen in both cell lines. The few shared DEGs are highlighted on the left. All genes are expressed as a Log_2 fold change vs PAR.

5.2.3 Proteome analysis of PAR and RAL Knockout cell lines using RPPA

Although, RNAseq data allowed us to comprehensively profile changes at the genome level, assessing changes on the level of protein expression is also required. RPPA not only determines protein expression levels but also phosphorylation, cleavage and fatty acid alteration modifications. Over 400 antibodies are utilised in this assay, that specifically recognise proteins in the PI3K-AKT and MAPK cascades and receptor tyrosine kinases, as well as components involved in cell cycle, DNA repair and apoptosis pathways. The same SW403 and SW620 clones (different lysates) used for RNAseq underwent RPPA analysis for comparison.

Again, the Pearson correlation coefficient of RPPA expression between each two samples within the same cell line was calculated (Figure 5.8A). Overall, the RPPA expression profiles of the RAL KO samples differ to the parental (PAR) samples in both SW403 and SW620. Whilst SW620 clone BKO2 was the most similar to the PAR samples, correlation of all KO clone trends in the same direction. All RPPA samples then underwent hierarchical clustering by the expression level of all protein and phospho-protein (Figure 5.8B). For SW403, RAL KO samples had expression profiles that were distinct from the parental cell lines. However, expression profiles varied between RALKO clones and were not isoform specific. The parental clones were also clustered together in the SW620s (Figure 5.8B), with only the SW620 clone BKO2 having a similar expression profile to the parental rather than the other RALKO samples.

Whilst a fold change cut off value for ±2.00 was used in the RNA sequencing analysis, very few proteins reached this level of expression difference and none were deemed to be significantly different following statistical analysis. Furthermore, RPPA analysis covers less of the proteome than the RNA sequencing analysis does of the transcriptome, therefore a higher cut off value may result in the loss of too many proteins/phospho-proteins. It was decided a ±0.5 cut off for defining a differentially expressed protein (DEP) would be more appropriate to reduce the risk of missing interesting and relevant proteins that could be investigated further at a later stage.

More differentially expressed proteins (DEPs) were identified in the SW620 RALKO samples than SW403 (Figure 5.9 and 5.10). The RPPA expression

values of the top RPPA hits in both cell lines, trend in the same direction as their corresponding gene expression obtained from the RNAseq data (Figure 5.9A and 5.10A). Only three DEPs in the SW403 had opposing gene expression (LDHA, PCNA and EphA2) (Figure 5.9A). This implies that expression trends at the gene level translates to the resulting protein, although it is important to note that RPPA expression values in individual clones may not be as strong as the other (Figure 5.9B and 5.10B).

As with the RNAseq data, the RPPA expression values for proteins and phospho-proteins trended in the same direction for both RALA and RALB KO samples in SW403s and SW620s (Figure 5.9 and 5.10). This supports the idea of the RAL isoforms carrying out similar functions. Only co-receptor CD4 and phospho-EphA2 had contrasting expression levels between RALKO status in the SW620s, with a decrease observed in the RALA KOs, and an increase in expression in the RALB KOs. Comparison of top proteins/phospho-proteins between cell types revealed very few trends in expression (Figure 5.11).

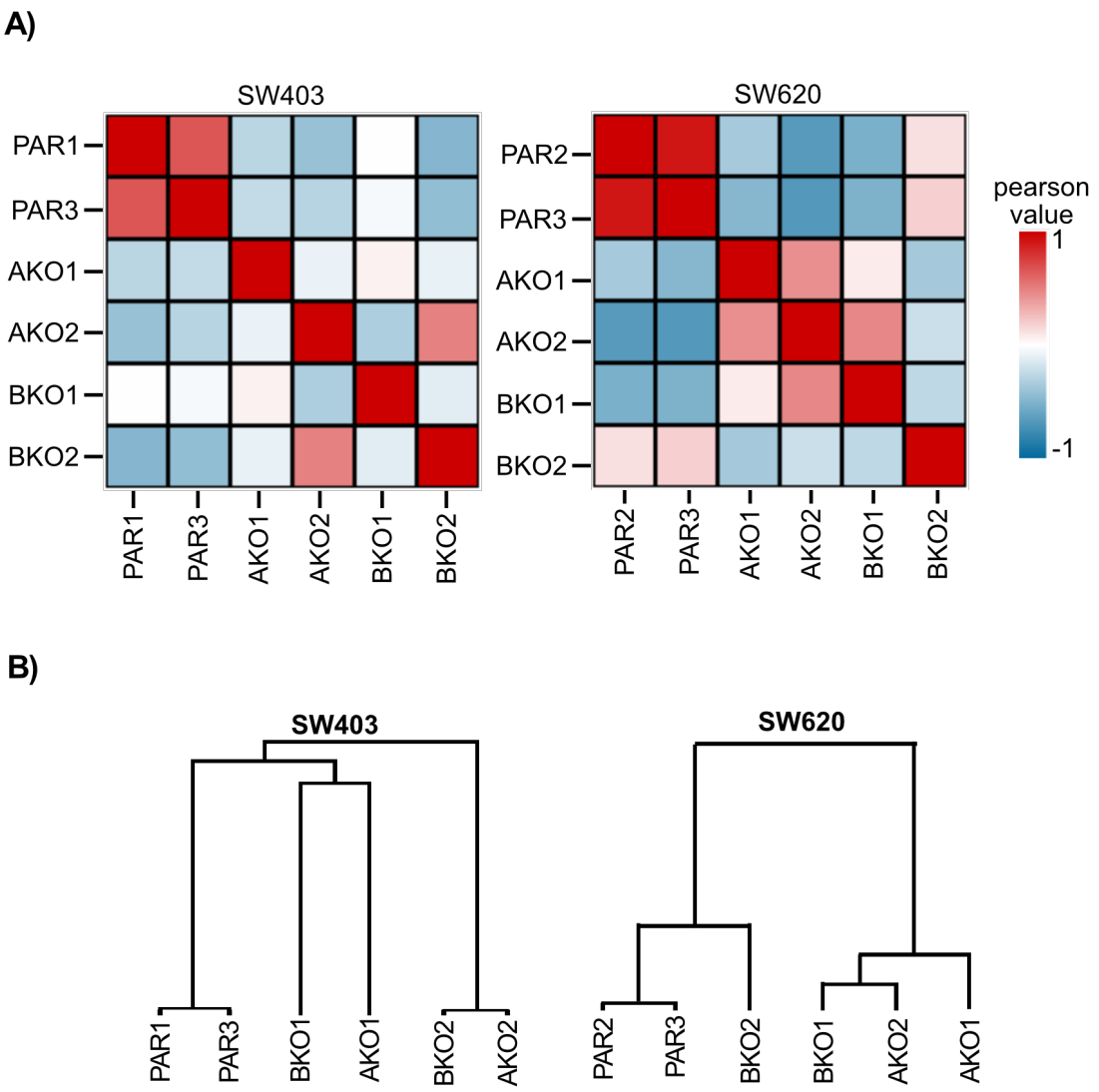


Figure 5.8 Negative correlation of protein expression between parental vs knockout clones in RPPA data analysis.
A) Pearson correlation coefficient values for all RPPA expression levels between each two samples. A negative correlation between all knockout clones (KO) and the parental (PAR) can be seen in both cell types. **B)** Hierarchical clustering of the expression level of all RPPA between each sample. All KO samples are displayed as distinct branches compares to the parental clones. Only SW620 clone BKO2 displayed the least change vs PAR in proteome expression compared to the rest of the knockout clones.

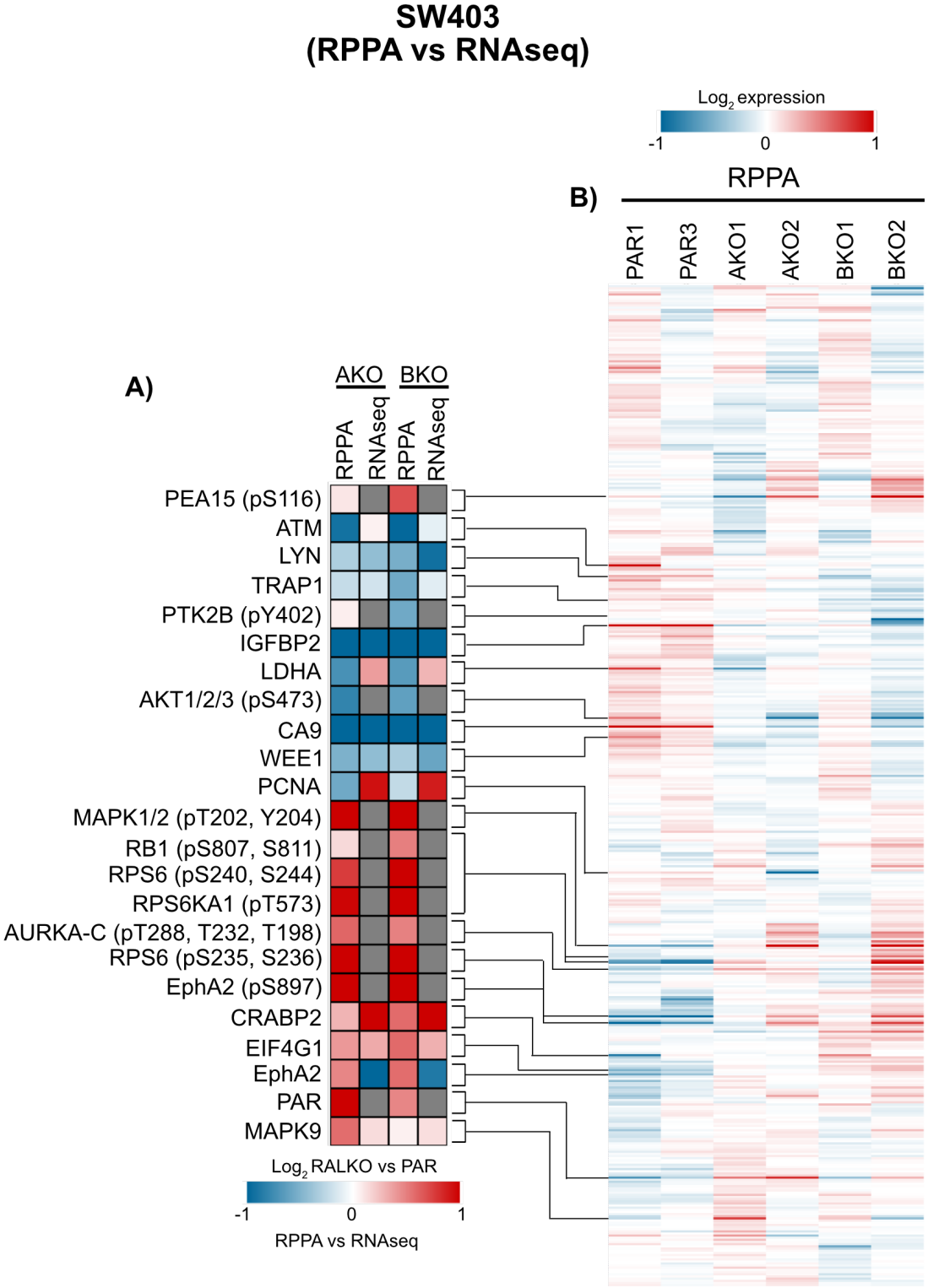


Figure 5.9 RNA expression matches top DEPs in SW403 when comparing RNAseq and RPPA data.
A) Differentially expressed (phospho-)proteins (DEPs) in the RAL KO vs PAR SW403 samples with the corresponding Log_2 RNA expression. **B)** Individual Log_2 expression values of all the RPPA data collected, with top hits (DEPs) highlighted. DEPs = Log_2 fold change ≥ 0.5.

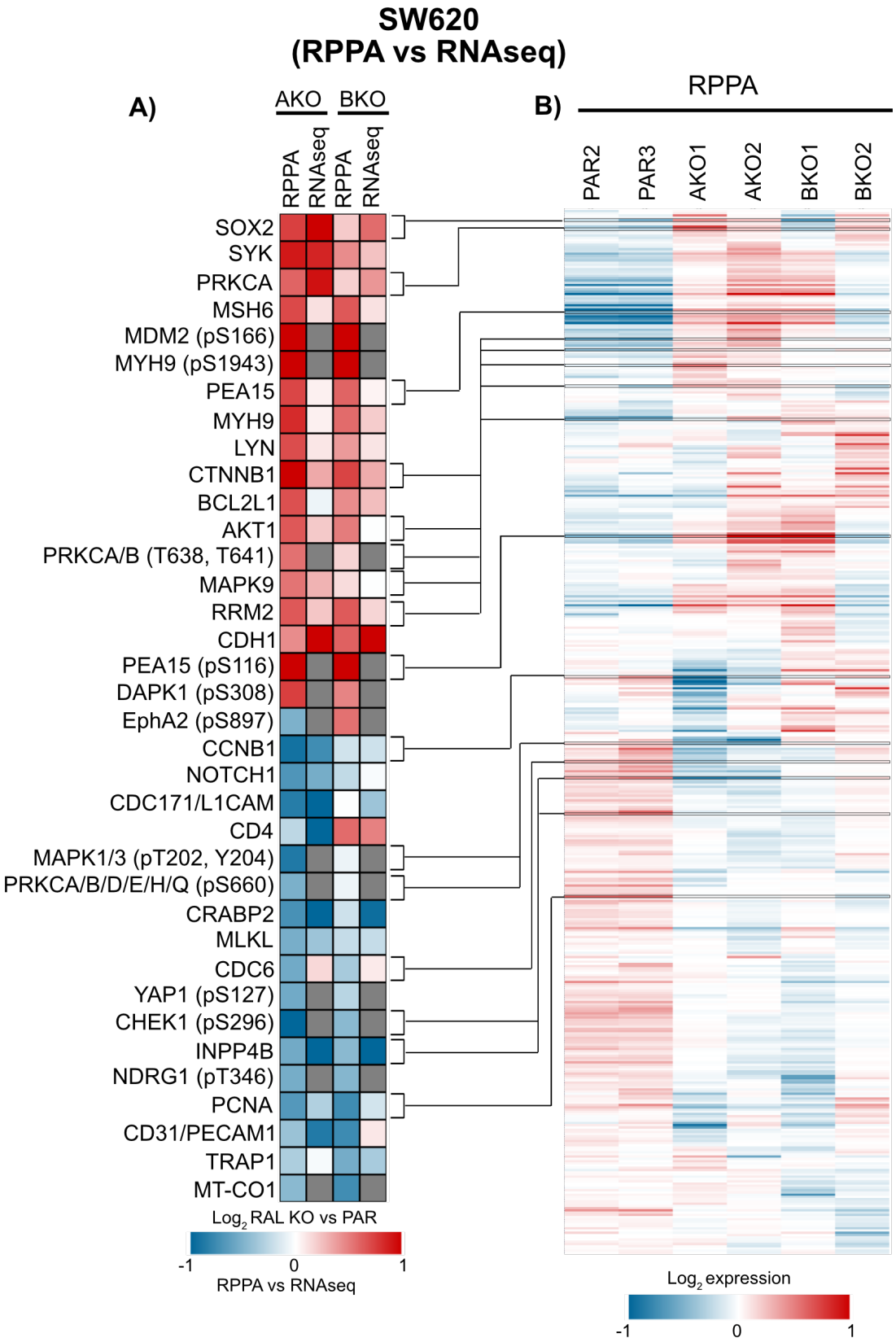


Figure 5.10 RNA expression matches top DEPs in SW620 when comparing RNAseq and RPPA data.
A) Differentially expressed (phospho-)proteins (DEPs) in the RAL KO vs PAR SW620 samples with the corresponding Log_2 RNA expression. **B)** Individual Log_2 expression values of all the RPPA data collected, with top hits (DEPs) highlighted. DEPs = Log_2 fold change ≥ 0.5.

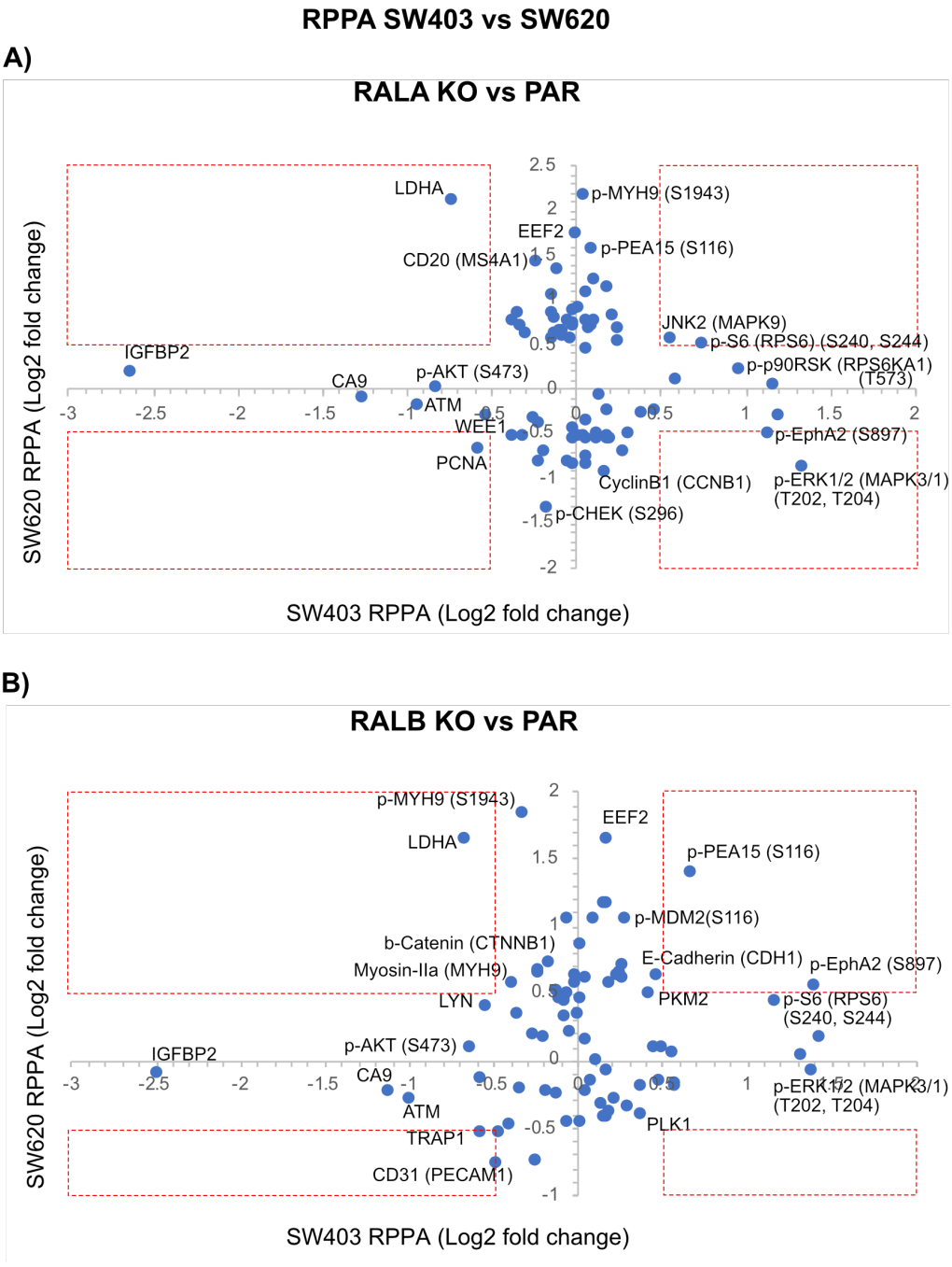


Figure 5.11 Comparison of top RPPA hits between SW403 and SW620 samples.
A) Comparison of top DEPs in RALA KOs vs PAR between SW403 (x axis) and SW620 (y axis) samples. **B)** Comparison of top DEPs in RALB KOs vs PAR between SW403 (x axis) and SW620 (y axis) samples. Differentially expressed (phospho-)proteins (DEPs) = Log_2 fold change ≥ 0.5. Red boxes highlight proteins that are classified as DEPs in both cell lines.

DEPs were separated into RALKO status vs PAR (Figure 5.11) and the majority of hits for one cell line did not have a significant expression fold change in the other or did not change at all. Once again, this highlights the overall different profiles of protein expression between cell lines that was also seen at the genome level. Only a few hits could be classified as DEPs for both SW403 and SW620 samples (highlighted in the red boxes), although expression did not always trend in the same direction between cell lines. Notable RPPA hits for both cell lines were the increased expression of JNK2 (MAPK9), phospho-S6 and phospho-PEA15, as well as increased and decreased expression of phospho-ERK1/2 in SW620s and SW403s respectively.

5.2.4 Pathway Trends and Gene Ontology (GO) Analysis

An overview of the RAS pathway (Figure 5.12) including key nodes across the RNAseq and RPPA data sets allows individual gene level with corresponding protein and phospho-protein expression values where available to be seen. In addition to this, GO analysis was carried out to gain an idea of the overall trends within both sets of data (Figure 5.13).

Gene Ontology (GO) analysis assigns enriched genes to predefined groups based on their functional characteristics, helping to build a more detailed gene expression profile and understand RAL isoform knockout effect on biological processes. Since the SW403 cell line had more DEGs compared with the SW620 in the RNAseq data set, GO analysis was performed on the SW403 DEGs and any enriched pathways were assessed for trends in the SW620 samples (Figure 5.13A). Individual genes that map to the processes highlighted in Figure 5.13A can be found in Supplementary Table 7.1.

Many different GO terms were associated with the DEGs provided, with the processes outlined in Figure 5.13 involving the largest groups of significant genes. The majority of processes are significantly decreased following either RALA or RALB knockout, with only DNA Replication, G1/S Transition of Mitotic Cell Cycle and Cell Mobility displaying an increase in related gene expression. Although these processes are presented as distinct functions, many of them involve crossover into different signalling pathways. The decrease in “EGFR

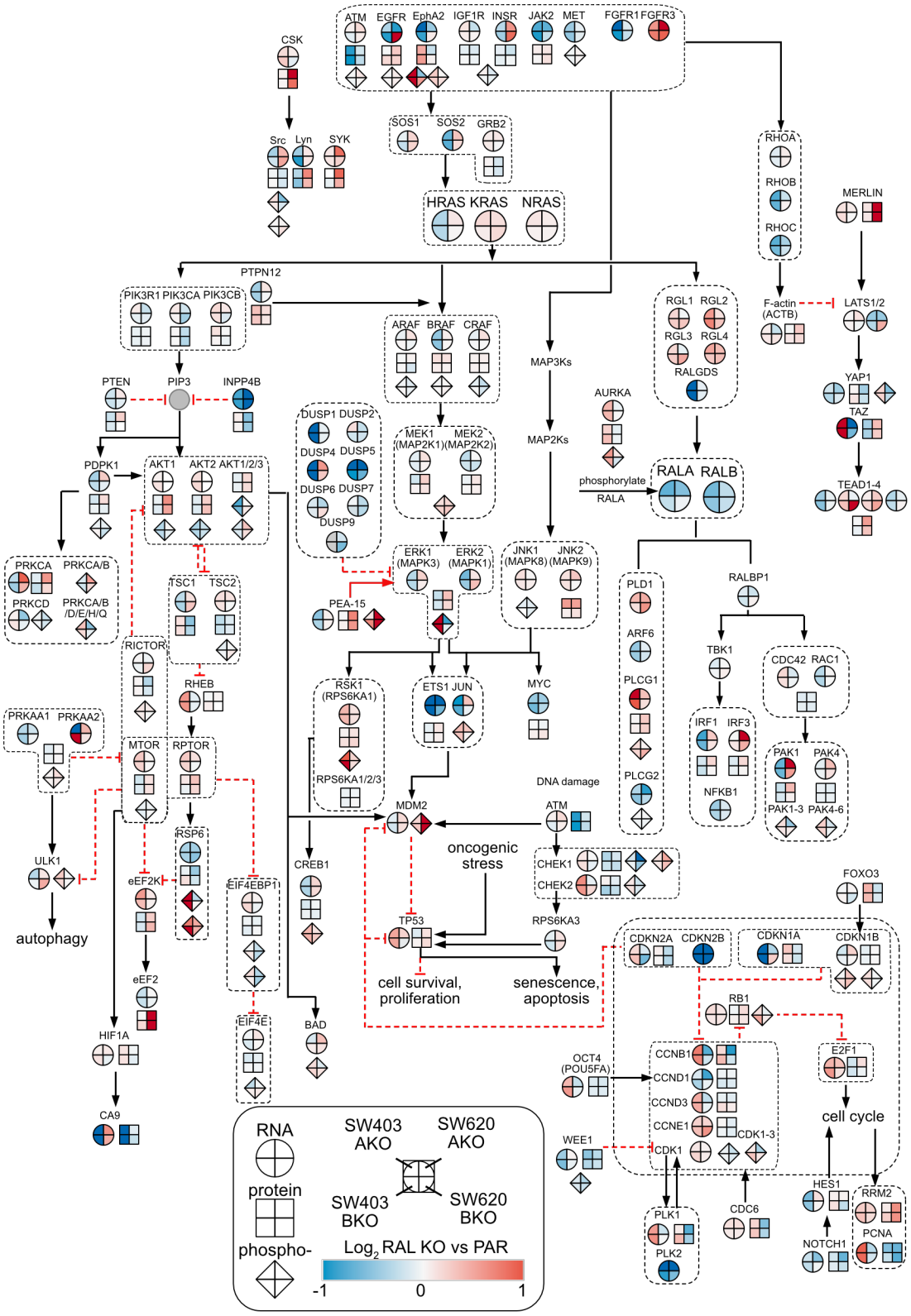


Figure 5.12 RAS signalling and cancer pathways.
An overview of the RAS pathways including the three main nodes; RAF-MAPK, PI3K-AKT and RALGEF-RAL pathways. Corresponding protein/phospho-protein expression is included from the RPPA data analysis where available. Expression is represented as a Log_2 fold change vs the parental (PAR).

Signalling Pathway" could impact Cell Proliferation and Apoptotic Processes, which themselves involve Transcription Factor Binding, Inactivation of the MAPK Pathway and Small GTPase Mediated Signal Transduction.

Many of the phenotypes identified in the GO analysis are well known to be driven by RAS signalling, although in many cases it is the MAPK/ERK and PI3K-AKT pathways that are generally considered to be the key regulators. The fact that they are enriched in the RAL KO cells suggests that there may be crosstalk with these other RAS effector pathways. The extent to which RAL isoforms might directly or indirectly regulate canonical RAS effector pathways is currently poorly understood and so this was selected as a line of investigation in our subsequent analysis.

Whilst the RNAseq data highlighted some interesting trends for future validation, changes in mRNA levels do not necessarily mean that there will be meaningful changes in protein levels of signalling wiring relevant to the targets of interest. To address this, protein and phospho-protein changes collected in the RPPA analysis were utilised and the corresponding genes also underwent GO analysis to determine if similar biological processes were affected (Figure 5.13B and 5.13C)

Together, the detailed RNAseq and RPPA data, along with the broader GO analysis, provides evidence of crosstalk between the RALGEF-RAL and MAPK pathways. The RAL regulators and effectors themselves, as well as possible RAL involvement in RAS-driven processes such as cell cycle are also interesting points that need to be looked into going forwards.

GO analysis

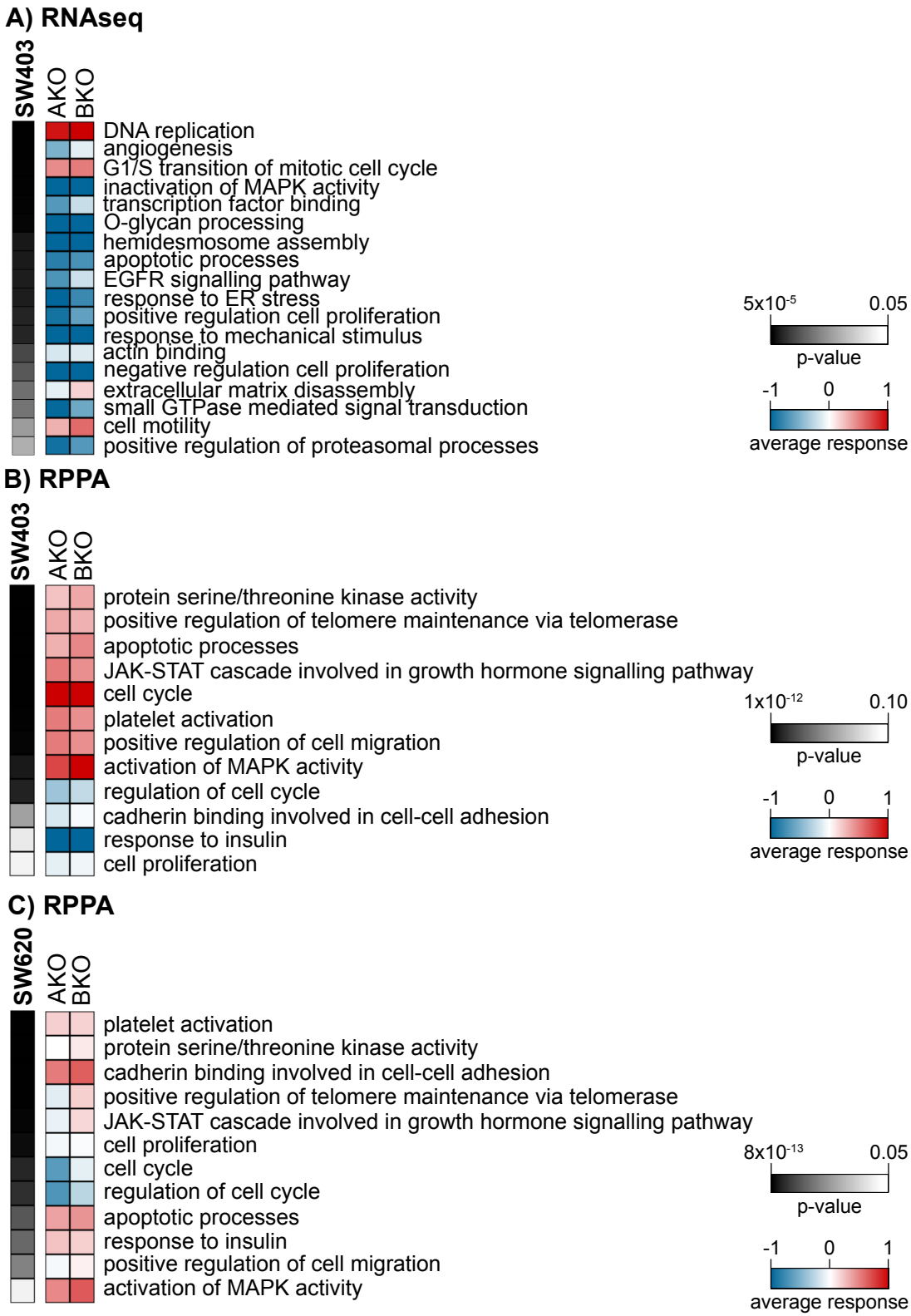


Figure 5.13 Gene Ontology (GO) Analysis highlights MAPK activity as well as proliferation and cell cycle pathways.
A) Go analysis of the SW403 DEGs collected from the RNAseq analysis. **B)** GO analysis of the top RPPA hits in SW403. **C)** GO analysis of the top RPPA hits in SW620. GO analysis for the RPPA data involved the genes of the top hits regardless of whether they were protein or phosphoproteins.

RAL signalling and cancer pathway trends – RNAseq and RPPA

RAL regulators/RAL pathway

RAL isoforms are regulated by a panel of GEFs and GAPs and in an isoform-specific manner by Aurora A kinases and Protein Kinase C. Once activated, RAL engages pathways regulating the exocyst, the actin cytoskeleton, phospholipase D. Surprisingly, analysis of the two datasets focusing on components of these known regulatory or effector functions, showed very few significant changes in gene expression (Figure 5.14A). In fact, only one DEG was identified in the SW403 cell line and that was a decrease in the RALGEF, RALGDS. Likewise, an increase in Interferon Regulatory Factor 3 (IRF3) gene expression in the SW620 RALA KO was the only DEG present in this cell line. Although few genes reached the high bar for differential expression there are some clear trends within the RAL regulators; RALGEFs and RALGAPs across the KO samples (Figure 5.14A). Even though no RPPA data is available for these RAL regulators, all four of the RAS-dependent RALGEFs RGL1-4 have increased expression in the RAL KOs compared to the parental, whilst RALGDS and the RAS-independent RALGEFs, RALGPS1 and RALGPS2 display a predominantly decreased expression. Similarly, all the RALGAP genes have decreased expression in the SW403 knockouts and are predominately increased in SW620s. Adding to these regulation changes, both RALA and RALB are also known to be phosphorylated by Aurora A kinase (AURKA) and Protein kinase C (PRKCA/B) respectively. Whilst RNA expression did not trend in the same direction between cell lines, it did translate to the protein and phospho-protein levels in both RALA and RALB KOs. However, this could not be confirmed by western blot (Figure 5.14B). Overall, knockout of one isoform seems to result in changes in RAL regulation but further validation is required.

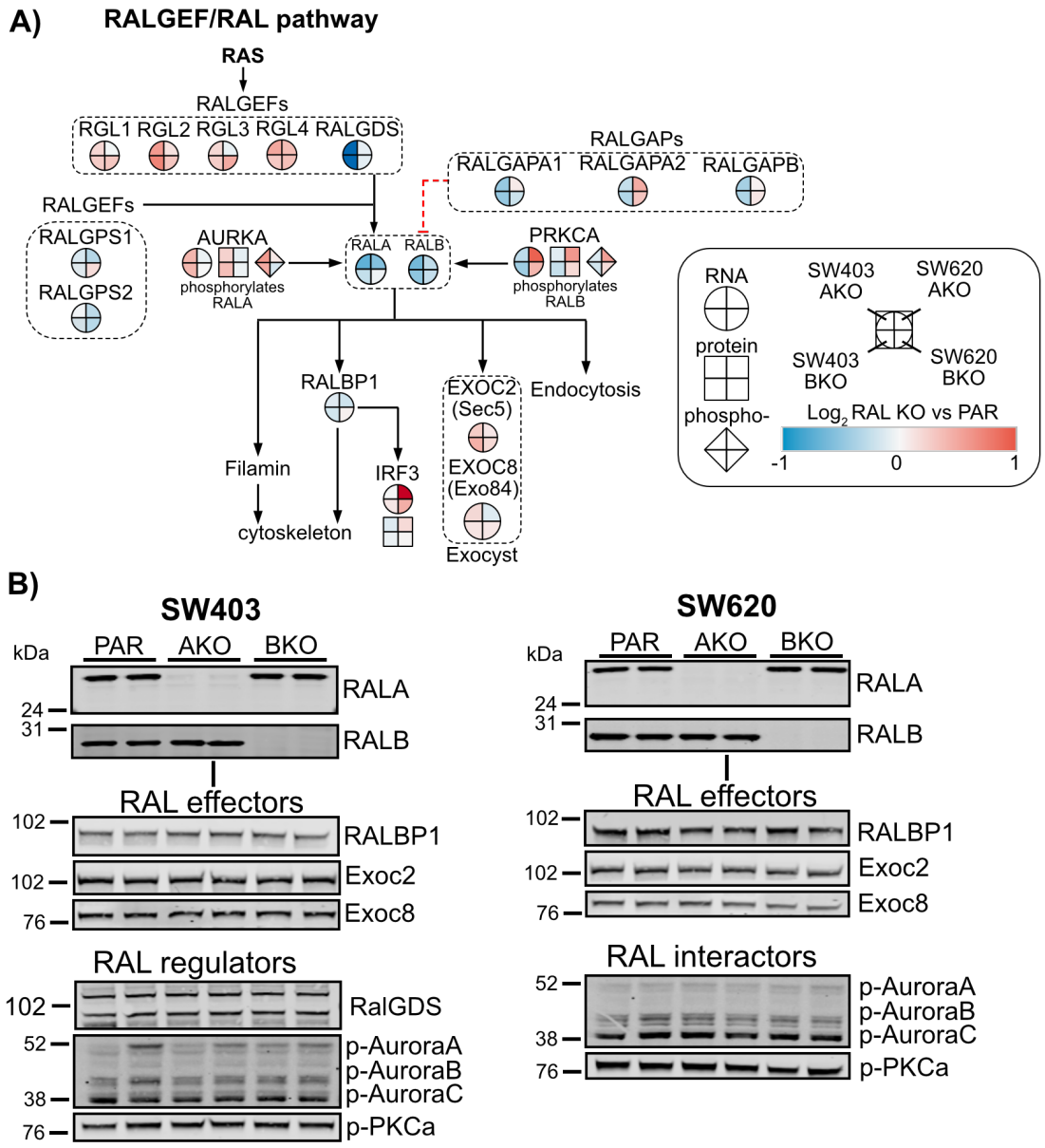


Figure 5.14 RNAseq and RPPA analysis of the RALGEF-RAL pathway.
A) An overview of the RALGEF-RAL pathway including RAL regulators and effectors. This includes components of known RAL functions such as the exocyst, filamin and cytoskeleton changes, endocytosis. Corresponding protein/phospho-protein expression is included from the RPPA data analysis where available. Expression is represented as a Log_2 fold change vs the parental (PAR). **B)** Validation of key RAL effectors and regulators by western blot revealed no observable changes in protein/phospho-protein expression, n=3.

RTK/MAPK pathways

Potential crosstalk with the other key RAS pathways was seen in many of the GO terms (Figure 5.13), including regulation of the MAP Kinases (MAPK). When looking at crosstalk, MAP Kinases involved in the JNK pathway were also considered.

The well-described MAPK/ERK pathway is centred on the activation of RAS and the subsequent activation of a kinase cascade involving RAF, MEK and finally ERK proteins. Knockout of either RAL isoform resulted in an overall gene expression decrease in the ERK regulators, DUSPs (Figure 5.15A). DUSPs are dual specificity phosphatases that dephosphorylate MAPK proteins. Therefore, a decrease in DUSP activity should lead to an increase in ERK activity. However, whilst the gene expression data collected suggests there may be a change in DUSPs, RPPA data was only available for DUSP4 and DUSP6 which didn't always translate from the mRNA level. As a result, DUSP5 expression was checked by immunoblotting with an increase in protein levels seen in the SW403 KO clones (Figure 5.15B). This is in contrast to the decrease in expression seen at the transcript level. Another but lesser known ERK regulator, PEA-15 is significantly increased at the protein and phospho-protein level in both the SW620 RALA and RALB KOs (Figure 5.15A). A smaller increase in phosphorylated PEA-15 is displayed in SW403, with a decrease at the gene and protein level.

Changes in ERK itself can be seen at the gene level (Figure 5.15A), right through to protein and phospho-protein. For the SW403 RAL KOs, there was a strong increase in phosphorylated ERK, yet the actual protein levels mimic those of the mRNA with a decrease in expression for RALA and RALB KOs. On the other hand, SW620s, have an increase in mRNA and total protein expression but a decrease in phospho-ERK1/2. However, just because the regulation of MAP kinases is different, doesn't mean the signalling of the pathway has also changed.

Although the gene expression trends differ between all the known ERK effectors, changes in these ERK related components suggest that there is evidence of MAP Kinase pathway regulation by RAL. Several transcription factors downstream of ERK were significantly decreased at the gene level. This included JUN (SW403), FOS (SW403 and SW620 RALA KO) and the

shared DEG, ETS1 highlighted in section 5.3.2. However only protein expression was available for ETS1 which actual displayed an overall opposite increase in expression and could not be validated by western blot (data not shown). Phosphorylation data was also obtained for JUN, which also contrasted to the gene expression, but it is important to remember mRNA status does not directly influence phospho-protein levels. The most consistent change in the MAPK/ERK downstream effectors was an increase in the ribosomal S6 kinase (RSK) family, particularly RSK1 an RSK3 (Figure 5.15A). These serine/threonine protein kinases are involved in cell growth, proliferation and survival. Similar gene expression data was observed for other MAPK pathways, including a decrease in DUSPs in both cell lines with a general increase in JNK MAP Kinases, but a lack of RPPA data covering these kinases and their effectors meant I was unable to perform a comprehensive analysis.

Finally, looking at receptors that are known to activate RAS, the receptor tyrosine kinase (RTK) EphA2 was found to be a DEG in SW403 RALA KOs (Figure 5.15A), as was EphA4 (data not shown). This translated to an increase in EphA2 protein and phospho-protein expression that was subsequently confirmed by western blot (Figure 5.15B and Supplementary Figure 7.12). However, re-expression of the knocked-out RAL isoform did not always restore p-EphA2 expression to that recorded in the parental (Figure 5.15B). Together, this RNAseq and RPPA data hints at both RAL isoform involvement in the regulation of all the MAPK pathways but these points need further validation to be understood properly.

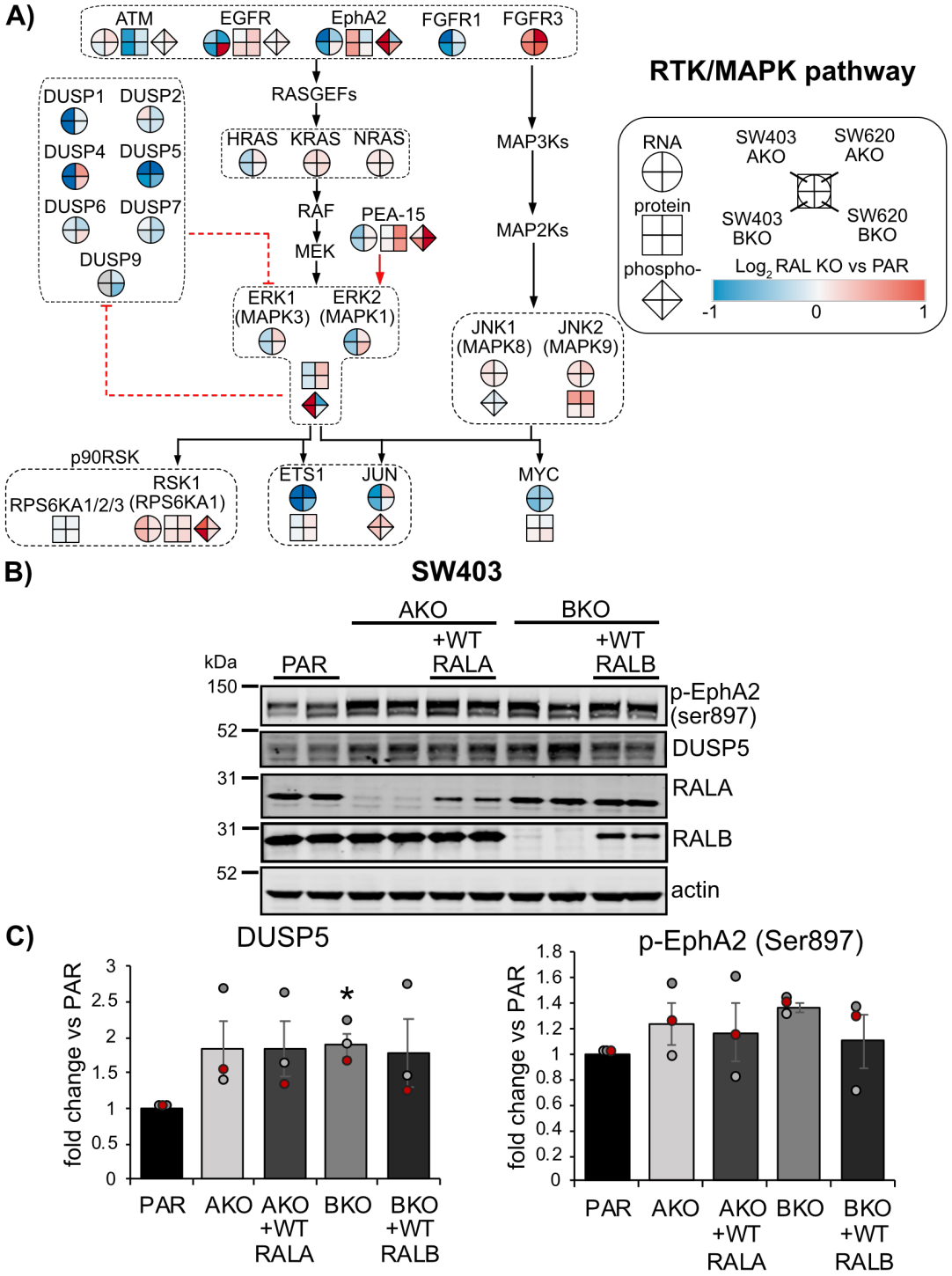


Figure 5.15 RNAseq and RPPA analysis of the RTK/MAPK pathway.
A) Overview of the RTK/MAPK pathway including receptor tyrosine kinases (RTKs). This pathway is predominantly involved in cell growth and proliferation. Corresponding protein/phospho-protein expression is included from the RPPA data analysis where available. Expression is represented as a Log_2 fold change vs the parental (PAR). **B)** Validation of RPPA hits phospho-EphA2 (ser897) and DUSP5 in SW403 PAR, RALKO and RALKO + WT RAL. An expression increase in both proteins can be seen in the RALA and RALB knockout clones (KO). Transduction of WT RALA/B back into KO clones did not restore expression levels back to parental. Although re-expression of RAL was not to the levels seen in the PAR clones. **C)** Quantification of DUSP5 and p-EphA2 expression. A significant (*) increase in DUSP5 expression is present in the BKO clones. Data is presented as a fold change and represents an increase or decrease in expression compared to PAR; fold change ± SE of n = 3 biological replicates.

Cell cycle

GO analysis highlighted cell cycle as being a significant process affected by knockout of RAL in both cell lines (Figure 5.13). The cell cycle is a 4-stage process consisting of a Gap 1 (G1), Synthesis (S), Gap 2 (G2) and mitosis (M) (Figure 5.16A). Each phase is regulated by the formation of cyclin (Cyclin A, B, D, E) and cyclin dependant kinases (CDK 1, 2, 4, 6) heterodimers (Figure 5.16A), which in turn are regulated by cyclin dependent kinase inhibitors (CDKN1A/B and CDKN2A/B/C).

Whilst gene expression changes between the different cyclins and CDKs is not consistent within each cell line, there appears to be an overall increase in related transcription factors and an overall decrease in cell cycle regulators (Figure 5.16B), albeit this is mostly centred on RNAseq. The most visible expression trends are in the various cell cycle regulators. The CAK complex genes (CCNH and CDK7) that activate the cyclin/CDK heterodimers are decreased in the RAL KOs compared to the parental in both cell lines, as well as cyclin B1 (CCNB1) and CDK1 negative regulator WEE1 (Figure 5.16B). The decrease in WEE1 has been validated by western blot in SW403 KOs and was found to be significant in both the RAL AKO and BKO clones (Figure 5.16D), although expression levels did not fully return to those of the PAR clones following re-expression of WT RALA or RALB (Figure 5.16D). There is also a decrease in cyclin B1 (CCNB1) and CDK1 regulator PLK1 in SW620 and an increase in SW403. Cyclin B1 itself, strongly reflects this expression pattern and is increased in SW403 and decreased in SW620 at the gene level right through to protein expression. The previously mentioned PEA-15 also plays a role in the cell cycle, regulating CDK6 and CDC25 directly. Both mirror the PEA-15 expression in each cell line, with an increase in the SW620 and a decrease in SW403, although RPPA data is absent for CDK6.

Like members of the MAPK pathways, analysis suggests that RAL has involvement in regulating the cell cycle. Whether this is through independent mechanisms or the previously mentioned MAPK pathways, which is heavily involved in cell cycle progression, needs to be looked into further.

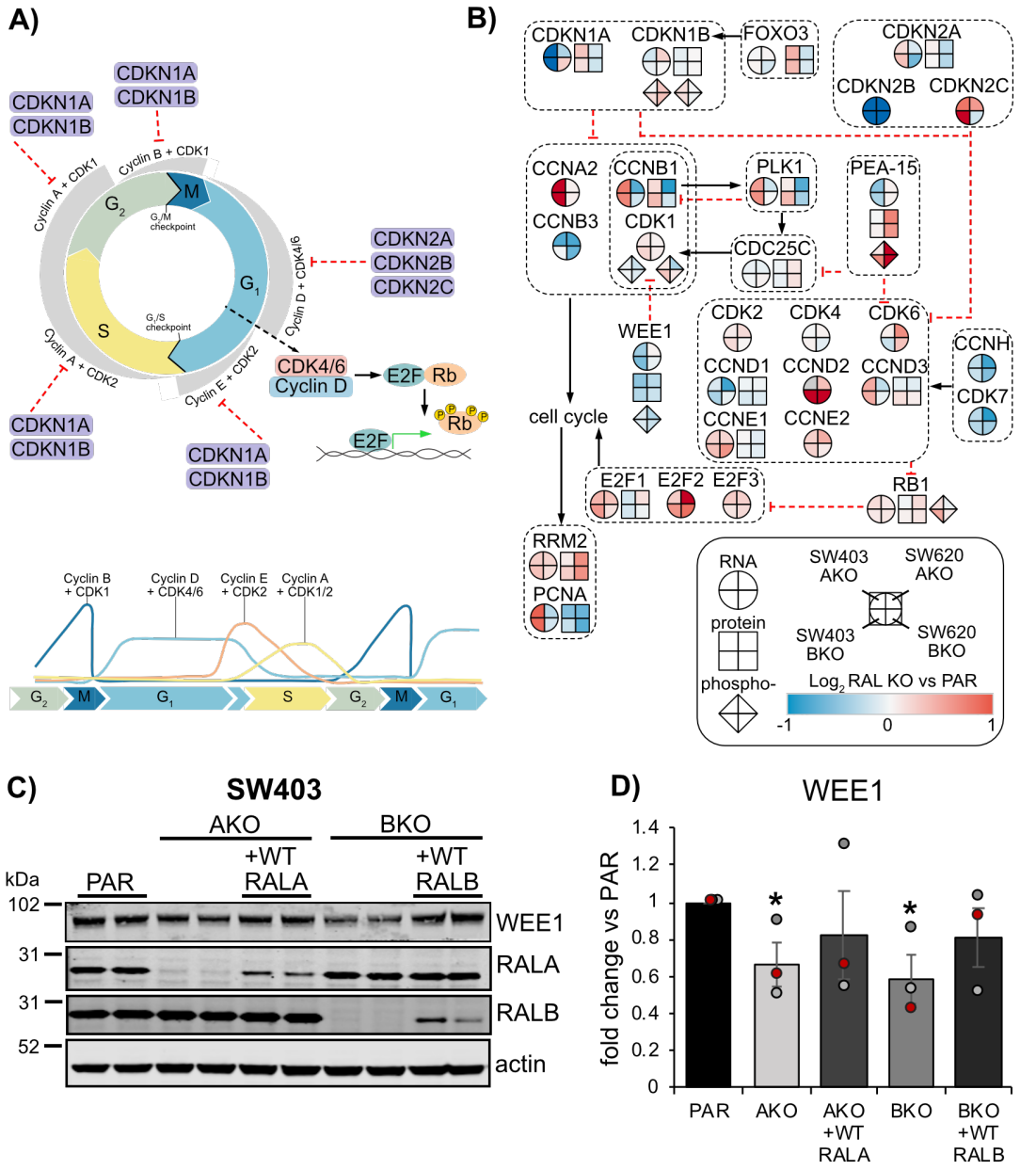


Figure 5.16 Changes in expression within the cell cycle.
A) The four distinct phases of the cell cycle; Gap 1 (G1), DNA synthesis (S), Gap 2 (G2) and Mitosis (M). Cell cycle regulation is carried out by the formation of cyclin (Cyclin A, B, D, E) and cyclin dependant kinases (CDK 1, 2, 4, 6) heterodimers. Cell cycle checkpoints are regulated by cyclin dependent kinase inhibitors (CDKN1A, CDKN1B). Cell cycle progression occurs from Cyclin D + CDK4/6 induced phosphorylation of Rb, which enables E2F-mediated transcription of S-phase genes. **B)** Cell cycle signalling network illustrating the fold change of RNA/protein/phospho-protein expression in SW403/SW620 RALA and RALB knockout cell lines. Red indicates a positive Log_2 fold change and an increase in expression compared to PAR (parental), whilst blue represents a negative fold change and a decrease in expression. **C)** Validation of RPPA hit WEE1 in SW403 PAR, RALKO and RALKO + WT RAL. A decrease in WEE1 protein expression can be seen in both RALA and RALB knockouts (KO). Transduction of WT RALA/B back into KO clones did not fully restore expression levels back to parental. Although re-expression of the RAL isoforms themselves was not to the endogenous levels seen in the PAR clones. **D)** Quantification of WEE1 expression. Quantification confirmed a significant (*) decrease in WEE1 expression ($p < 0.05$) in both AKO and BKO vs PAR. Data is presented as a fold change and represents an increase or decrease in expression compared to PAR; fold change ± SE of n = 3 biological replicates.

5.3 Discussion

In this chapter I report the first large-scale analysis of RAL GTPases in a colorectal cancer cell model using a multi-omic approach. This involved profiling the network of outputs influenced by each RAL isoform using genome-wide and protein expression/activity analysis. Integrating the isoform activity and multi-omic data has revealed the unique RAL signalling signatures of different CRC cells.

When it comes to changes in RAL isoform activity, previous work has mainly focused on activation as a consequence of a non-RAL loss/gain of function. The activity of RALA has been found to increase significantly in PI3Kγ knockout mice (Sim *et al.*, 2013), knockdown of RASGAP RASA1 leads to an increase in RALA activity via R-RAS (Sung *et al.*, 2016) and RAB10 increases the GTP binding of RALA through the recruitment of Rgl2 (Karunanithi *et al.*, 2014).

Here, the effects on activity of each RAL isoform following gene knockout of RAL itself have been determined. Our data shows that upon deletion of one RAL ortholog, the activity levels of the remaining isoform increase, whilst total protein levels remain similar to the parental. This was observed across all four cell lines with a significant increase in RALB activity detected in all cell lines except SW48. This could suggest that the mutant RAS proteins present in the HCT116, SW403 and SW620 cell lines impact RAL activity, as the WT RAS in the SW48 cell lines does not increase RAL activity to significant levels. The idea that there is compensatory activation of one RAL isoform in the absence of the other was highlighted in isoform specific null MEFs (Peschard *et al.*, 2012). Activity levels of GTP-RALB were found to be significantly increased in RALA null MEFs and GTP-RALA was increased in RALB null MEFs. The increased activation was also reported to occur without any apparent changes in RAL protein levels. This data from a mouse model study supported these findings and in fact, a significant increase in GTP-RALA in the BKOs, in one of the cell lines investigated was found.

The reasons for how this compensatory increase occur remain unclear but it is not unreasonable to believe this is due to the removal of competition through

an increase in availability of RALGEFs and RALGAPs. A deficiency of RALGAP, RALGAPA1 has been found to cause constitutive activation of RALA (Wagner *et al.*, 2020), RALGAP suppression caused RALB activation and the suppression of autophagy (Martin *et al.*, 2014) and aberrant overexpression of the RALGEF Rgl2, promotes RAL-dependent pancreatic cancer growth (Vigil *et al.*, 2010). Therefore, if RAL regulator availability directly affects RAL activity, it is not unlikely that the availability of RAL itself, say the lack of one isoform, can affect the activity levels of the other.

Why compensation occurs is likely linked with the functional redundancy of the two isoforms. It also highlights the question of RAL isoform abundance and how much of each is present that consequently determines the level to which one isoform can compensate for the other. Likewise, the different RAL regulator abundances in different cell lines will also impact the levels of RAL activity compensation and therefore RAL-dependent functions. The idea of RAL isoform abundance will be addressed in the next chapter but varying RAL quantities across cell lines will ultimately influence RAL function and dependency on this particular pathway.

The idea that one isoform compensates for the other is supported by a highly similar response at the mRNA level for RALA and RALB in both cell lines tested. Only a few genes showed opposite expression changes between RAL status upholding the premise of similar effector interactions. This could also explain why known effectors and interactors of the RAL pathway appear to show little to no change at the mRNA and protein level.

Despite the same trend response between RALA and RALB KOs for most genes within a cell line, there was very little correlation between the expression profiles across the two cell lines at the transcript level. Since both SW403 and SW620 cell lines harbour G12V KRAS mutations, it was thought that their RAL signalling responses, at least in terms of RAS activity, might be more comparable. This highlights how different the signalling networks between cell lines can be and that their genetic background can be extremely varied regardless of the fact they have the same tissue origin. This is likely due to variations and mutations in other signalling pathway proteins. One example of

this is the fact that SW620 is homozygous for mutant KRAS and the SW403 cell line only has a single mutated copy.

With this in mind, it was surprising that the SW403 KOs had a higher number of changes at the transcript level when compared to the parental, rather than the SW620 KOs. Studies have shown that increased copy number of mutant KRAS is present in a significant portion of human tumours (Soh *et al.*, 2009). This suggests that multiple copies of mutant KRAS may impart a stronger signalling effect than a single mutant copy. However when it comes to KRAS mutant colorectal cell lines and their dependency on RAL, it has been previously shown that knockdown of both isoforms decreased cell viability in the SW403 cell line more than SW620 with survival of the remaining population being ~50% and 75% respectively (Lee *et al.*, 2019). Therefore, the increased dependency on RAL for cell viability in the SW403 cell line, seems to be unrelated to mutant KRAS copy number and therefore may not contribute to the larger response seen in our RNAseq data following RALA or RALB knockout.

Only twelve genes were identified as being differentially expressed in both cell lines and only seven of those genes trended in the same direction (PLK2, MYZAP, EREG, GJB3, ETS1, PRSS3 and CHST3). Studies focusing on RAL and these genes is severely limited. Only a link between RAL and ETS Like-1 protein (ELK1) has been reported, in which growth hormone stimulated RALA-PLD formation was required for ERK activation and ELK1 transcription (Zhu, Ling and Lobie, 2002). Additionally, although displaying contrasting expression profiles between SW403 and SW620, IGFBP2 could potentially be a gene of interest. The IGFBP2 counterpart IGFBP1 was found to activate ERK and RAL in an oligodendrocyte precursor cell line (Chesik *et al.*, 2010). It's worth noting that IGFBP2 was also identified from the RPPA data as being differentially expressed at the protein level in SW403 KOs. In spite of the lack of research, several of these genes are either involved in MAPK pathways (EREG and ETS1) (Tetsu and McCormick, 2017; Kohsaka *et al.*, 2014; Draper *et al.*, 2003) or cellular processes linked to RAS signalling (PLK2) (Lee, Hoe and Pak, 2011; Lee *et al.*, 2011; Walkup *et al.*, 2018). Together with the RAL-dependent activation of ELK1, this data supports the idea of crossover between RAL and the other main pathways of RAS.

Unfortunately, apart from IGFBP2 and ETS1, these genes were not covered in the RPPA analysis to see if expression changes at the mRNA level translated to the protein level. Furthermore, western blot analysis of EREG, ETS1 and PLK2 expression did not match that seen in the RNAseq data. This lack of translational efficiency could be explained by modifications such as methylation of the mRNA affecting their stability (Mauer *et al.*, 2017; Zaccara, Ries and Jaffrey, 2019). The protein stability itself may also be a factor, as proteins have different half-lives (Eden *et al.*, 2011) with some proteins needing to be quickly degraded for proper function (Toyama and Hetzer, 2013). Indeed, the ETS1 ortholog ETS2 is known to have a half-life of approximately 20 minutes (Papas *et al.*, 1990), which is considered a very short half-life with the majority of proteins lasting between 8 and 10 hours (Chen, Smeekens and Wu, 2016).

Like the RNAseq, the RPPA data displayed different protein expression profiles between the SW403 and SW620 cell lines. However, there was still a high correlation of expression change between RALA and RALB, again reinforcing the idea that isoform compensation does not just involve activity levels but extends to functionality. As previously mentioned, there were very few DEGs that matched up with proteins that were differentially expressed, however clear trends could be seen throughout the two datasets supporting the change from mRNA to protein.

By comparing transcriptomic, proteomic and phospho-proteomic datasets, likely protein expression/activity changes and the potential participation of RAL in novel signalling pathways were identified. A high correlation between mRNA and protein levels indicated that protein abundance was likely to be determined by transcript abundance and that the RPPA dataset was reporting differential protein expression across both RAL knockout cell lines.

For this reason, focus was turned to proteins involved in the RTK/MAPK and PI3K pathways, as well as processes highlighted in the GO analysis that are known to be activated by RAS, such as cell proliferation and growth including the cell cycle. Although multi-omic analysis was only carried out on two out of the four knockout cell lines generated, potential hits were also validated in the HCT116 an SW48 cell lines where appropriate.

Unfortunately, some of these proteins that were highlighted by the RPPA and had corresponding RNAseq data, could not be validated by western blot. Probing for phospho-RSK, PEA15, phospho-PEA15 and phospho-ERK did not display any significant expression changes. Whilst this does mean further validation is required to establish a link with RAL, it should be noted that RPPA is a more robust form of quantification and has a higher degree of sensitivity compared to western blot (Boellner and Becker, 2015). This could explain why small changes in protein expression were detected in RPPA but not western blot.

From our initial activity studies, activation of both RAL isoforms was found on average to peak around one minute, extending to two minutes in HCT116. Whether or not the duration of the remaining RAL isoform also increases following isoform specific knockout has not been determined, it seems unlikely that it would extend far beyond those seen in the parental cell lines. With this in mind, any involvement of RAL in the activity of these proteins could be short-lived and provide the initial activation before other pathways become more dominant and any differences in expression are lost.

Finally, rewiring and compensation in the RAS pathways has been shown to occur in a matter of days. Phospho-ERK levels returned to control levels after only 24 hours treatment with a G12C-specific KRAS inhibitor (Hallin *et al.*, 2020), and work conducted in our lab has shown drug resistant signalling pathways emerge after only 2-4 days with KRAS inhibitor treatment (Warren *et al*, unpublished). Therefore, the expression changes seen in both the RNAseq and RPPA may have already been adjusted back to parental levels due to alternate pathway activation. Carrying out western blot validation on knock-in cell lines or even double RAL knockout might help address this issue by either enhancing protein interaction and activity changes via a constitutively active RAL or preventing potential compensation through the remaining RAL isoform.

Our studies provide evidence of a link between several DUSPs and both RAL isoforms. There was a clear downward trend of nearly all the DUSP mRNA levels in both the SW403 and SW620 cell lines, with DUSP4 and DUSP5 being significantly downregulated. Only data covering DUSP4 protein levels was

available from the RPPA and this was seen to have an inverse relationship with mRNA expression with a small increase in protein levels detected in the SW403 KOs.

Western blot analysis confirmed an increase in the nuclear protein DUSP5 in SW403 RALA and RALB KOs which again, is opposite to that reported at the transcript level. It should be noted that validation of a link between RAL and DUSP5 was not always successful, as this DUSP5 expression increase in SW403 KOs was not consistently restored to levels seen in the parental clones, following re-expression of the knocked-out isoform. This is most likely due to the re-expression levels of RAL itself being noticeably lower than that of the endogenous RAL. Future re-expression attempts, whether that be stable or transient expression, would benefit from having a stronger expression promoter, such as a CMV promoter, present in the transfer plasmid. Moreover, isolation and expansion of a positive re-expression cell population would also ensure RAL levels were as close as possible to those recorded in the parental clones and eliminate interference from any remaining knockout cells present in a heterogenous population.

Further validation efforts through RALA or RALB knockdown in the parental clones (data in supplementary material), did not show an increase in DUSP4 or DUSP5 expression. However, it is likely that the knockdown period was not sufficient for RAL to exert any effect. Additionally, knockdown did not fully eliminate isoform expression and therefore may also explain the lack of differential DUSP protein expression.

The MAPK pathway is subject to a large number of negative feedback loops that includes the dephosphorylation and inactivation by DUSPs (Caunt and Keyse, 2013). DUSP4 and DUSP5 display nuclear localisation and have a preference for different MAPK proteins, with DUSP4 inactivating ERK, p38 and JNK, whereas DUSP5 is ERK specific (Lake, Corrêa and Müller, 2016). Upon MAPK pathway stimulation, DUSP expression levels increase. Moreover, DUSPs can be directly phosphorylated by MAPKs, leading to increased stabilisation and prolonged activation, which in turn inactivates MAPK proteins. However, studying DUSP activity and MAPK regulation is difficult due to the presence of this feedback mechanism, short protein half-lives and low basal expression levels. Indeed, a DUSP2 deficiency was found to not cause

enhanced ERK phosphorylation (Jeffrey *et al.*, 2006). Similarly, DUSP10 knockdown does not induce p38 hyperphosphorylation (Zhang *et al.*, 2004).This inconsistency of DUSP regulation of MAPK proteins may be attributed to the use of overexpression systems in an effort to combat the low basal levels of DUSP expression.

By measuring the endogenous expression of DUSP5 mRNA levels after MEK inhibitor treatment, the molecular mechanisms underlying the induction of DUSP5 transcription by the MAPK pathway were finally revealed. These involved promoter binding sites for a ternary complex involving ETS binding sites for transcription factor ELK1, serum response factor (SRF) and a serum response element, with subsequent depletion of SRF or ELK1 leading to a decrease in DUSP5 mRNA following serum stimulation (Buffet *et al.*, 2015). This study also found that the mRNA of DUSP5 is rapidly induced by ERK activation but unlike DUSP6, its stability is independent from the activation of the MAPK pathway.

There is currently no available data connecting RAL activity and DUSP proteins. However, considering the link between RAL and ELK1 (Zhu, Ling and Lobie, 2002), DUSP5 and ELK1, as well as transcription factor ETS1 presenting in our data as a DEG in both SW403 and SW620, crosstalk between RAL and the MAPK pathway through DUSPs looks promising. Since DUSP5 stability seems unaffected by the MEK/ERK pathway, taking mRNA measurements by RT-qPCR following the combined use of a MAPK and PI3K inhibitor would assist in determining the potential role RAL plays in DUSP5 regulation. This inhibitor approach would also help deduce other interactions between RAL and RAS pathway kinases, such as p-ERK, RSK and AKT, that we were unable to fully validate by western blot.

One kinase that displayed evidence of RAL dependency was receptor tyrosine kinase EphA2. In the SW403 RAL KO cells, a downregulation of EphA2 mRNA and an increase in protein and phospho-protein expression was observed, which was subsequently confirmed by western blot. EphA2 is generally thought of as a tumour promoter and is overexpressed in a variety of different cancers including and colorectal cancer (Kou and Kandpal, 2018). EphA2 displays significantly higher expression levels in colorectal tumours leading to

an increase in cell migration and invasion and correlates with poor survival in stage II/III cancer tissues (Dunne *et al.*, 2016). Importantly, this study found that EphA2 levels were regulated by KRAS-driven RALGDS-RALA and MAPK pathways. Further analysis revealed that RNAi-mediated silencing of RALA but not RALB significantly decreased EphA2 protein levels in HCT116 and DLD-1 cells. Furthermore, examination of a large colorectal cancer cell panel revealed a link between high EphA2 expression and an activating mutation in KRAS. Like DUSPs, a negative feedback loop exists between RAS and EphA2. It has been shown that whilst RAS-MAPK signalling induces EphA2 expression, EphA2 stimulation through ligand interaction, downregulates MAPK signalling (Macrae *et al.*, 2005). This may explain why the EphA2 mRNA and protein expression levels reported in the SW403 cells are differentially expressed.

Whilst a link between RAL and EphA2 has already been established, the data reports the opposite expression profiles to those seen in our data. This could simply be due to different cell lines used or the fact that the association between the two proteins was detected as a result of RNAi not full gene knockout. Since our data has demonstrated a compensatory effect in RAL isoform activity, not only does this suggest that EphA2 regulation may be a PAN-RAL response, but our knockout cells have had a longer time to adjust and alter this pathway. Regardless, collectively these data provide evidence that EphA2 is an important mediator in colorectal cancer progression and that this RAL-dependent function may be a consequence of oncogenic RAS.

Finally, several cell cycle proteins were identified from our two datasets as being differentially expressed in both cell types including an overall downregulation of CyclinD1 and WEE1 mRNA and protein levels, as well as changes in CyclinB1 expression that translates from mRNA to protein. RAS activity is known to promote cellular proliferation or differentiation through growth factor stimulation and ultimate activation of MAPK proteins. MAPK induced phosphorylation of transcription factors such as FOS, JUN and ETS (identified as DEGs in our data) results in the expression of immediate-early response genes required for progression through early phases of the cell cycle (Zhang and Liu, 2002). Whilst RAL has not been shown to contribute to cell proliferation in the absence of RAS (Drosten *et al.*, 2010), RAL has been

implicated in cell cycle progression. Notably, RAL GTPases have been found to enhance cell detachment and motility through Cyclin D1 (Fernández *et al.*, 2011) and the regulation of Cyclin D1 by RAL occurs through activation of NFκB (Henry *et al.*, 2000). One study has also shown that activation of RAS during the mid-G1 phase of the cell cycle does not result in ERK activation and that RAS may target an alternative effector pathway during cell cycle progression (Taylor and Shalloway, 1996). Could RAL be responsible for the passage through multiple points during the cell cycle? Indeed, WEE1 is a protein kinase that regulates the G2 checkpoint and prevents entry into mitosis by inhibiting phosphorylation of CyclinB1/Cdc2 complexes (Geenen and Schellens, 2017). Activation of RAL by RAS has been found to interfere with this phosphorylation process and decrease CyclinB1/Cdc2 activity (Agapova *et al.*, 2004). The reduced expression of WEE1 collected from the omic analysis was validated by western blot, providing evidence for a link between RAL and cell cycle progression through WEE1 regulation.

The combination of techniques was applied to obtain an insight into protein interactions and signalling pathway adaptations to isoform specific RAL knockout. Importantly it highlights how varied the signalling networks of different cell lines can be and that an appropriate cell line with robust dependency on the target protein pathway, making it capable of highlighting novel interactions, is crucial in determining its role in cellular physiology. Together, these data provide evidence for crosstalk between RAL and other pathways of RAS as well as highlights involvement with novel regulators of known RAS-driven processes.

Chapter 6: Protein Standard Absolute Quantification of RAL Isoforms

6.1 Introduction

Despite studies spanning over three decades, a complete understanding of RAL GTPase function and dysfunction, particularly in cancer, is yet to be realised. Indeed, one part of RAL biology that has not been fully addressed is protein abundance. Specifically, how much RAL, as well as how many molecules of each RAL isoform are present in cell lines commonly used in cell signalling cancer research. Protein abundance is a critical parameter in cell signalling, especially since research is becoming less focused on identifying and characterising individual proteins and instead shifting to a more holistic approach that seeks to understand how proteins work together in complex signalling networks (Aldridge *et al.*, 2006; Bordbar *et al.*, 2014). Therefore, utilising a method that can serve to quantify proteins of interest is highly desirable.

Although the utilisation of immunoblotting to quantify small GTPase protein abundance and activation is a well-established technique (Omerovic *et al.*, 2008; Chapman and Hall, 2005; Kim *et al.*, 1997), in this instance it fails to provide a larger, more detailed picture of RAL biology that is required. Whilst it would enable the identification of RALA and RALB and their expression in cells or tissue samples, immunoblotting is highly variable and would not allow sufficient quantification and therefore comparison between isoforms. Furthermore, due to the lack of PAN RAL antibody availability, the overall RAL expression across a panel of cell lines would not be accurately determined. Therefore, a more robust method to measure the abundance of RAL isoforms is preferred.

The emergence of Liquid Chromatography with tandem mass spectrometry (LC-MS/MS)-based proteomics has provided a more effective method for studying complex proteomes. This approach involves the enzymatic digestion of a proteome to generate peptides that are subsequently acidified to give them a positive charge before separation through LC and analysis by MS. Due to its ability to identify thousands of peptides/proteins, this techniques is often

referred to as shotgun proteomics (Aebersold and Mann, 2003; Domon and Aebersold, 2006). Unfortunately this approach has limited sensitivity and most often only provides relative quantification of the most abundant components (Gallien, Duriez and Domon, 2011).

Thus, alternative targeted proteomic approaches such as selected reaction monitoring (SRM)-based LC-MS have been developed to quantitatively and precisely analyse complex samples (Anderson and Hunter, 2006). This assay usually involves the initial ionisation of a sample followed by two stages of mass selection: first selecting the mass of the intact "parent" ion, and then that of a specific fragment ion following collision-induced dissociation. Collectively, this generates a SRM assay, or multiple reaction monitoring (MRM) if more than one parent/fragment ion pair is monitored (Gallien, Duriez and Domon, 2011). Furthermore, this MS-based approach in combination with an appropriate isotope-labelled standard can provide absolute quantitation of endogenous protein concentration.

Since RAL has currently not been quantified in this way but both RAL isoforms share similar (46-51%) sequence identity and domain architecture to that of RAS small GTPases (Nicely *et al.*, 2004), studies that have attempted to quantify cellular RAS abundance using this technique were looked into. By adding isotopic RAS reference peptides as standards ahead of SRM analysis, a number of groups directly measured RAS protein abundance but with highly variable results, ranging from <100 to >5,000,000 molecules of RAS per cell (Schwanhäusser *et al.*, 2011; Halvey, Ferrone and Liebler, 2012; Wang *et al.*, 2011).

The significant divergence in RAS abundance estimates recorded in these studies is likely due to the use of enrichment steps and relative quantitation strategies. All these targeted approaches involved the labelled RAS peptide standards being spiked-in immediately prior to injection of the sample into the mass spectrometer. This means that any differences acquired during sample preparation, proteolysis and/or peptide gel extraction were not factored into the final quantitation (Lowenthal *et al.*, 2014). To avoid these problems, Mageean et al. optimised a methodology known as Protein Standard Absolute Quantitation (PSAQ) (Mageean *et al.*, 2015). Instead of adding individual peptides, PSAQ involves the addition of full-length isotope-labelled protein

standards into the sample cell lysate before processing and importantly does not require any antibody enrichment steps (Brun *et al.*, 2009; Dupuis *et al.*, 2008). By adding the stable recombinant protein standard at the beginning of the procedure, it will be subjected to the same potential errors and losses as the endogenous protein, enabling a more accurate approach to protein quantification.

Mass spectrometry analysis performed by SRM-based quantification, in combination with a protein standard, paves the way for quantifying endogenous levels of specific proteins and their isoforms. Importantly, because expression of both RAL isoforms is not a result of ectopic over-expression, data across any cell line of interest can be accurately collected.

6.1.1 Aims and Objectives

The aim of this chapter is to precisely quantify RAL isoform abundance across a large panel of colorectal cell lines. The objectives of this process involved the establishment of the PSAQ method in the lab. This required the following stages of optimisation:

- Produce isotope-labelled His-tagged recombinant RAL proteins.
- Optimise sample processing through in-gel digest.
- Identify and validate proteotypic RAL peptides and mass spectrometry transitions.

6.2 Results

6.2.1 Production of PSAQ RAL standards

As mentioned previously, the generation of full-length, isotope-labelled protein standards forms the basis of the PSAQ strategy and is required for the successful quantification of a target protein(s) (Figure 6.1). Since the proteotypic RAL peptides selected later on will contain a C-terminal lysine residue, His-tagged RALA and RALB recombinant proteins were produced labelled with $^{13}C_6$ $^{15}N_2$ L-lysine (+8 Da) and $^{13}C_6$ $^{15}N_4$ L-arginine (+10 Da). To express these proteins, AT713 bacteria deficient for the biosynthetic genes ArgA and LysA were used, since they are suitable for generating uniformly labelled proteins with arginine and lysine isotopes.

Following sequence cloning of WT RALA and RALB into the pTrcHis A vector, AT713 E. *coli* were transformed with these plasmids and grown under different conditions in M9 minimal media which had been supplemented with heavy arginine and lysine. Culture samples taken and run on SDS-PAGE revealed His-RALA was stable at 37°C and 21°C but an overnight incubation at the lower temperature increased the yield. On the contrary, His-RALB seemed unstable when grown overnight, even at 21°C. Therefore, standards for RALB

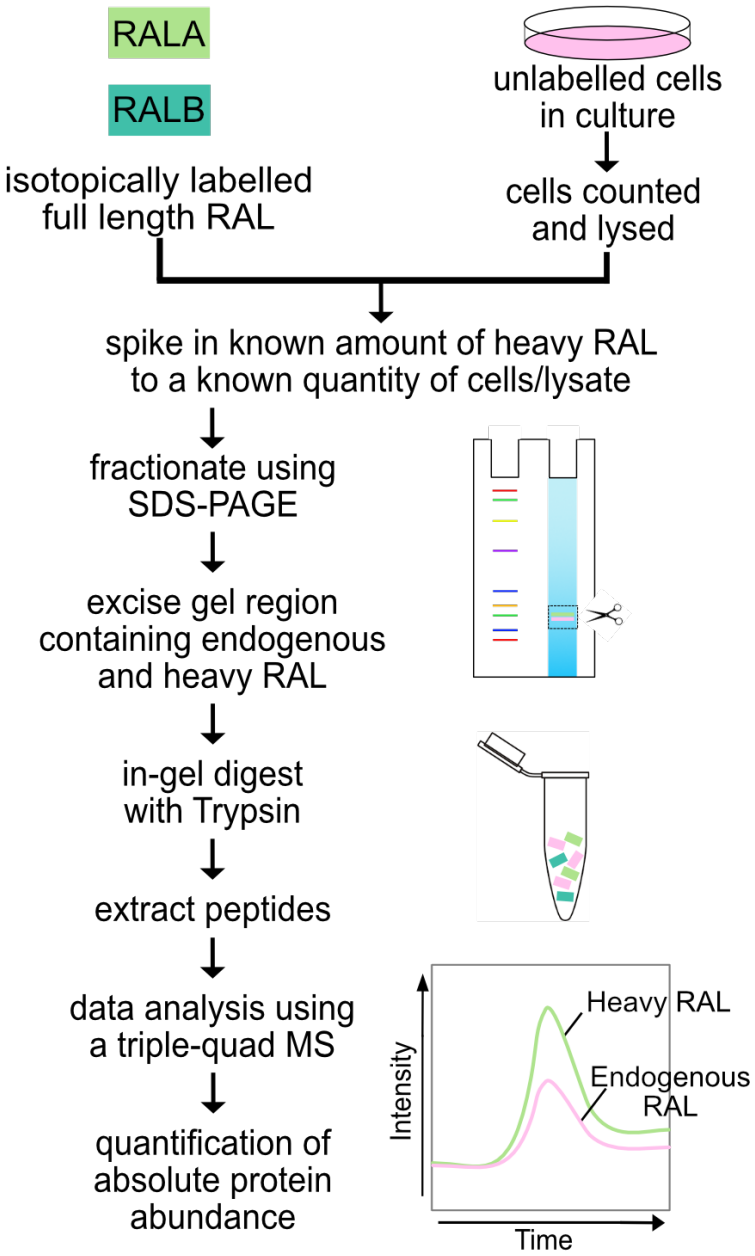


Figure 6.1 Protein Standard Absolute Quantification (PSAQ) for measurement of cellular RAL abundance.
Isotopically labelled RALA and RALB recombinant proteins are spiked into cell lysates at a known concentration before fractionation, proteolysis and quantification of pre-determined proteotypic peptides using selected reaction monitoring to allow calculation of RAL isoform abundance.

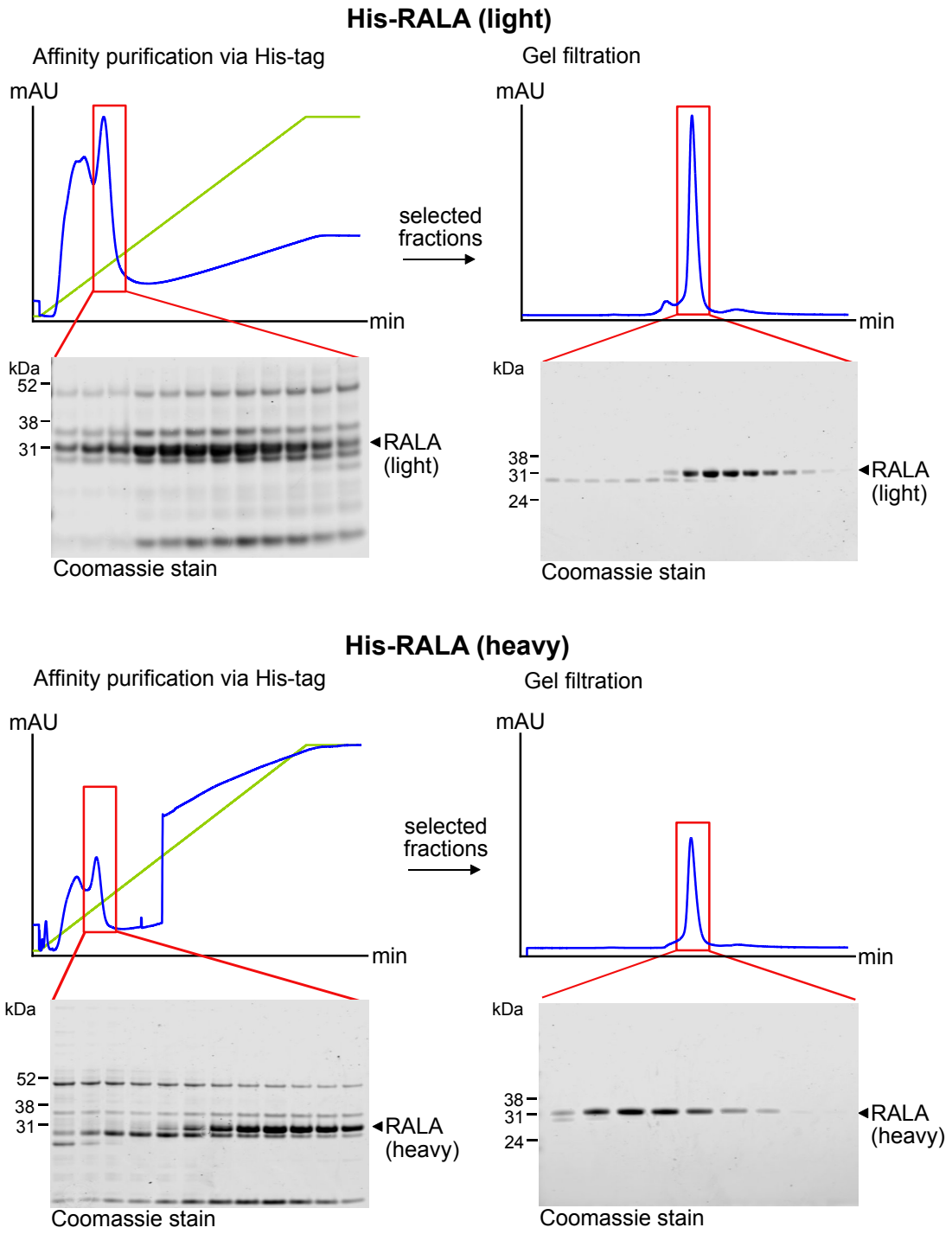


Figure 6.2 Elution profiles for His-tag and size exclusion-based chromatography of RALA (light) and RALA (heavy) purifications.
Nickel-based enrichment of His-tagged proteins typically retains contaminating proteins that require further purification steps to achieve sufficient purity for use as an isotope-labelled protein standard for use in protein quantitation. Blue = UV absorbance (mAU), Green = percentage concentration of Buffer B (%).

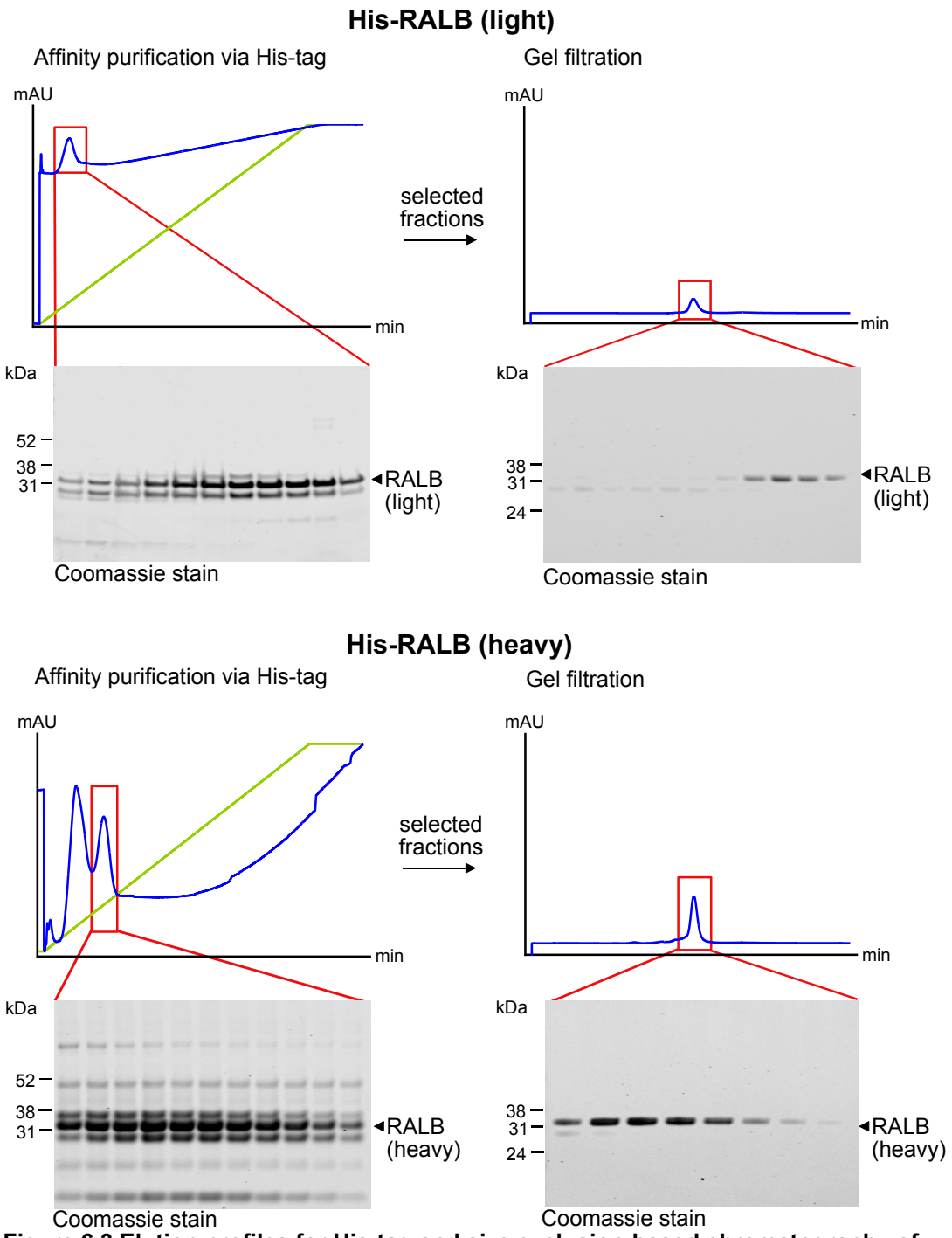
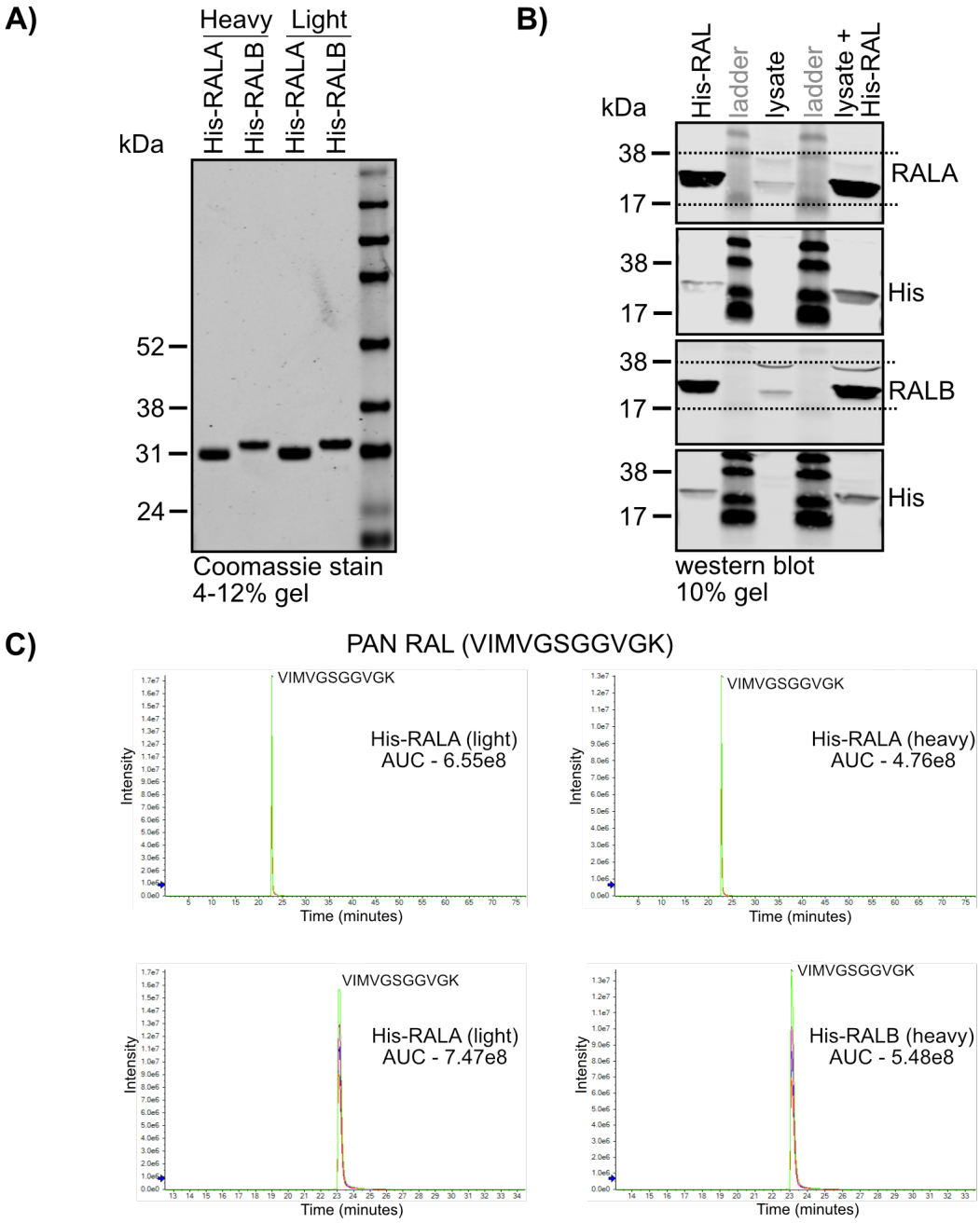


Figure 6.3 Elution profiles for His-tag and size exclusion-based chromatography of RALB (light) and RALB (heavy) purifications.
Nickel-based enrichment of His-tagged proteins typically retains contaminating proteins that require further purification steps to achieve sufficient purity for use as an isotope-labelled protein standard for use in protein quantitation. Blue = UV absorbance (mAU), Green = percentage concentration of Buffer B (%).

were produced at 37°C with only a 3-hour incubation to prevent possible degradation and His-tag cleavage (Supplementary Figure 7.15). The His-tagged RAL isoforms were subsequently purified using His-Trap HP columns before relevant fractions were collected and subjected to gel filtration using Superdex Increase 200 10/300 GL and AKTA purifier equipped with Frac950 and UPC900 (Figure 6.2 & Figure 6.3). Further details of purification are in the supplementary material (Supplementary Figure 7.16-7.19).

Following this last step of purification, coomassie staining of fractions containing the RAL PSAQ standards revealed there was no contaminations with any other proteins (Figure 6.4A). The concentration of the RAL PSAQ standards was determined by generating an isotopically light labelled His-RALA protein. The concentration of this light version was measured using a bicinchoninic acid (BCA) assay (Figure 6.4D). Once an accurate concentration was determined, a known amount of the light His-RALA was mixed 1:1 with the heavy RAL PSAQ standards, before in-gel digestion and SRM analysis of the light and heavy versions of the PAN RAL peptides (Figure 6.4C). Accurate measurement of the PSAQ standards was resolved through comparison of the intensities of the respective isotope signals.

Finally, 15 µg of SW48 cell lysate with and without 100 ng of PSAQ standards was run on a self-poured 15 well, 1 mm, 10% BisTris SDS-PAGE gel with alternating lanes of rainbow marker to determine the migration of the endogenous and His-tagged RAL proteins and which segment is required for processing (Figure 6.4B).



D)

			concentration		
isoform	peptide	sequence	BCA mg/ml	actual mg/ml	ng/µl
RALA	Pan RAL	VIMVGSGGVGK	0.230	0.167	167
RALB	Pan RAL	VIMVGSGGVGK	0.384	0.282	282

Figure 6.4 High-purity His-RAL protein standards.
A) Coomassie staining showed no contaminating proteins. 200 ng of each protein loaded per lane. **B)** Determining gel migration for accurate excision of target protein and standards. The His-tag results in a band shift that must be accounted for when determining where to cut the gel. For the western blots, 5 ng His-RAL and 15 µg SW48 lysate are loaded in the indicated lanes. Note all recombinant proteins run at different band sizes depending on gel composition. **C)** A known amount of accurately quantified His-RALA light (100 ng) was mixed BCA reported 100 ng worth of isotope-labelled RAL standards to accurately determine heavy-labelled protein concentration. **D)** Concentration of His-RAL proteins determined by BCA and then by mass spectrometry analysis.

6.2.2 Identification of proteotypic RAL peptides and their transitions

To be able to accurately quantify protein abundance, it is first necessary to identify peptides that are unique to the protein(s) of interest. Identifying these peptides, known as proteotypic peptides (Mallick *et al.*, 2007), can present a significant challenge with small GTPases such as RAL, as the two isoforms share an 82% overall sequence identity (Gentry *et al.*, 2014). This limits the number of isoform specific peptides that can be generated following proteolysis as a high proportion of those peptides generated will be shared between both isoforms.

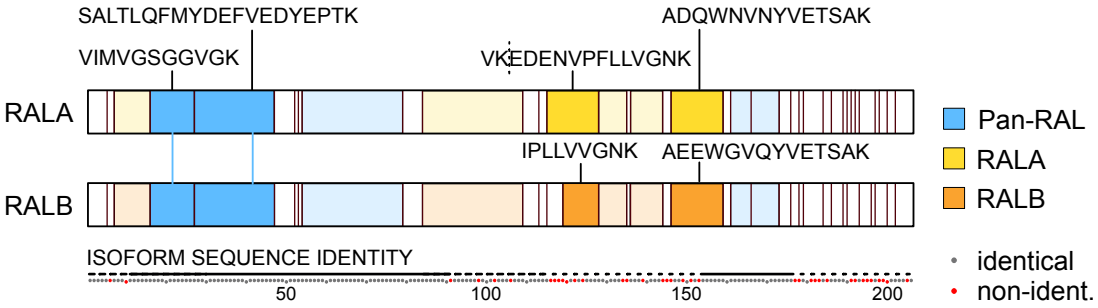
That said, only a few representative peptides are needed to confirm the presence of a protein in a biological sample (Lange *et al.*, 2008). However, due to the considerable amount of time and resources required to produce quantitative standards, such as these recombinant proteins, the selection of the best peptides for each protein is a crucial step in the development of SRM assays (Deutsch, Lam and Aebersold, 2008). Therefore, peptides were carefully selected and the following criteria were used to decide which ones were most suitable to measure RAL abundance: (1) peptides must be unique to both RAL isoforms or isoform specific, (2) ideally peptides must be between 8-25 amino acids in length, (3) a peptide must have limited chemically modifiable residues, such as asparagine, glutamine and methionine (4) a peptide must be devoid of known post-translational modification sites e.g. not be from the HRV region of RAL, (5) exclusion of peptides generated as a result of missed cleavages is preferred and (6) peptides with multiple mid-chain basic residues should be discounted.

The majority of large-scale proteomic studies utilise trypsin to digest proteins into analysable peptides (Gallien, Duriez and Domon, 2011). A similar protein quantification study carried out in our laboratory, found that trypsin generated the most information regarding RAS abundance from a single digestion. Trypsin is a stable protease that displays high cleavage specificity at arginine (R) and lysine (K) residues and can act under a wide range of conditions (Olsen and Mann, 2004). Since trypsin cuts after a K or R, it means that the peptides should always be labelled with a heavy isotope, either 8 or 10 Da respectively. The presence of a C-terminal basic residue in tryptic peptides

means they are typically multiply charged, making them highly suitable for fragmentation under low energy collision induced dissociation (CID) conditions (Paizs and Suhai, 2001). For these reasons, trypsin was selected as the enzyme of choice for digestion of the samples collected.

Tryptic digestion of both RAL isoforms was first performed *in silico* using the Expasy PeptideMass calculator. To ensure specificity to RAL, all selected peptides were subjected to BLAST analysis (Basic Local Alignment Search Tool, National Library of Medicine) using a non-redundant protein sequences database and blastp (protein-protein BLAST) algorithm. The predicted peptides that agreed with most of the criteria outlined above, were selected as potential candidates (Figure 6.5). Although, it should be noted that it was difficult to find any PAN RAL or isoform specific peptides that did not include at least one modifiable residue.

To discover if these peptides were detectable under our experimental conditions, the recombinant proteins underwent in-gel digestion and analysis on the mass spectrometer at different concentrations with and without cell lysate. After selection of suitable transitions (Figure 6.6 & Figure 6.7), initial analysis of a high concentration (100 ng) of either His-RALA/B alone (without cell lysate) included detection of all candidate peptides (Supplementary Figure 7.20 and Supplementary Figure 7.21). However, the levels of RALB specific peptide IPLLVVGNK were found to be lower than that of the others and RALA specific peptide EDENVPFLLVGNK was found to include a miscleavage (VKEDENVPFLLVGNK).



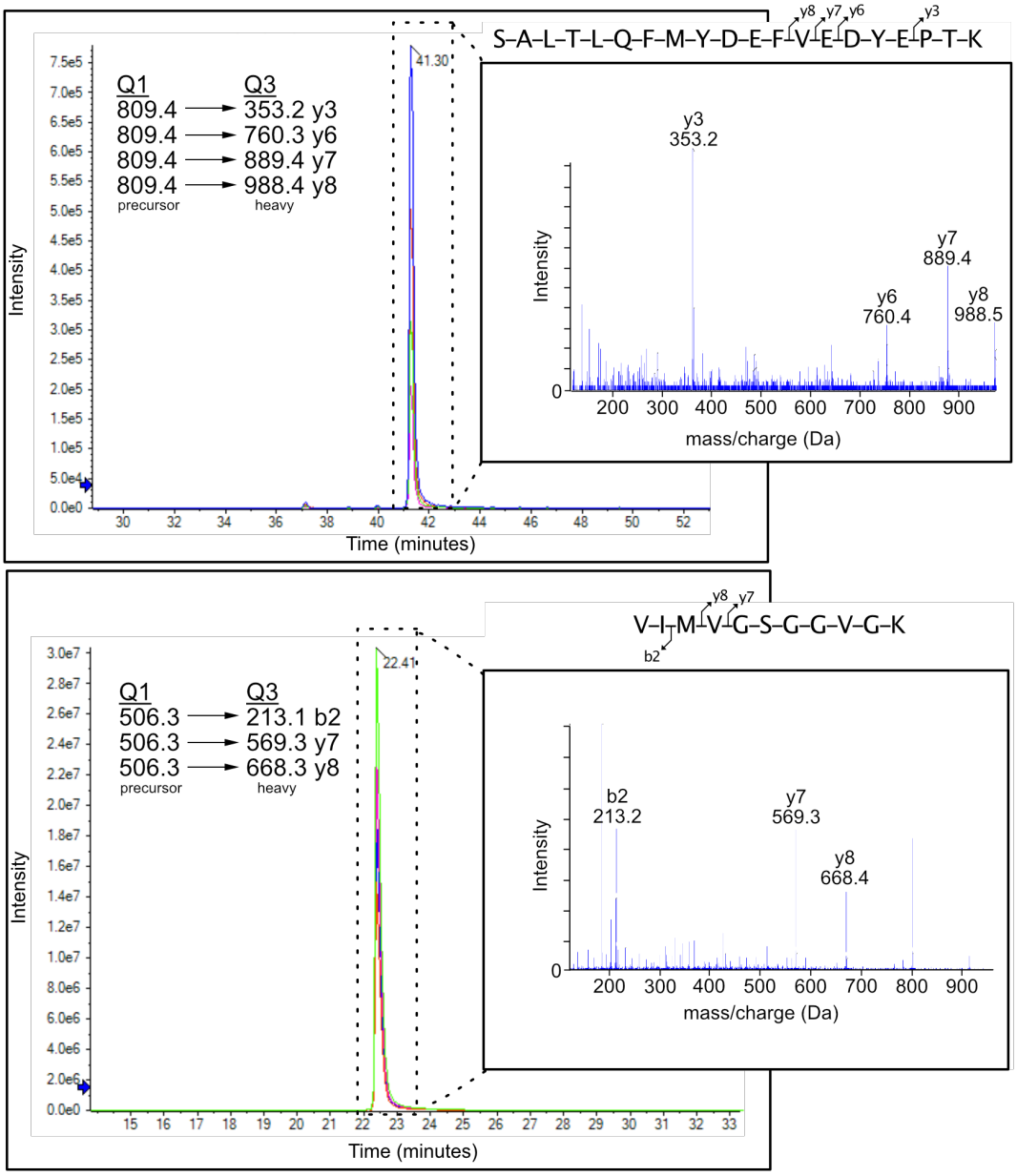
RAL Isoform	PEPTIDE	sequence	a.a positions
RALA	Pan RAL	SALTLQFMYDEFVEDYEPTK	28-47
	Pan RAL	VIMVGSGGVGK	17-27
	RALA	AEQWNVNYVETSAK	146-159
	RALA	VKEDENVPFLLVGNK	114-128
RALB	Pan RAL	SALTLQFMYDEFVEDYEPTK	28-47
	Pan RAL	VIMVGSGGVGK	17-27
	RALB	AEEWGVQYVETSAK	147-160
	RALB	IPLLVVGNK	121-129

Figure 6.5 Proteotypic RAL peptides.
All peptides are tryptic-derived. The candidate peptides describe total RAL as well as isoform specific peptides. Only RALA specific peptide EDENVPFLLVGNK was consistently detected with a miss-cleavage (dashed line) in practice runs.

This missed cleavage site is likely due to the presence of a glutamic acid (E) at the start of the desired peptide sequence (Šlechtová *et al.*, 2015).

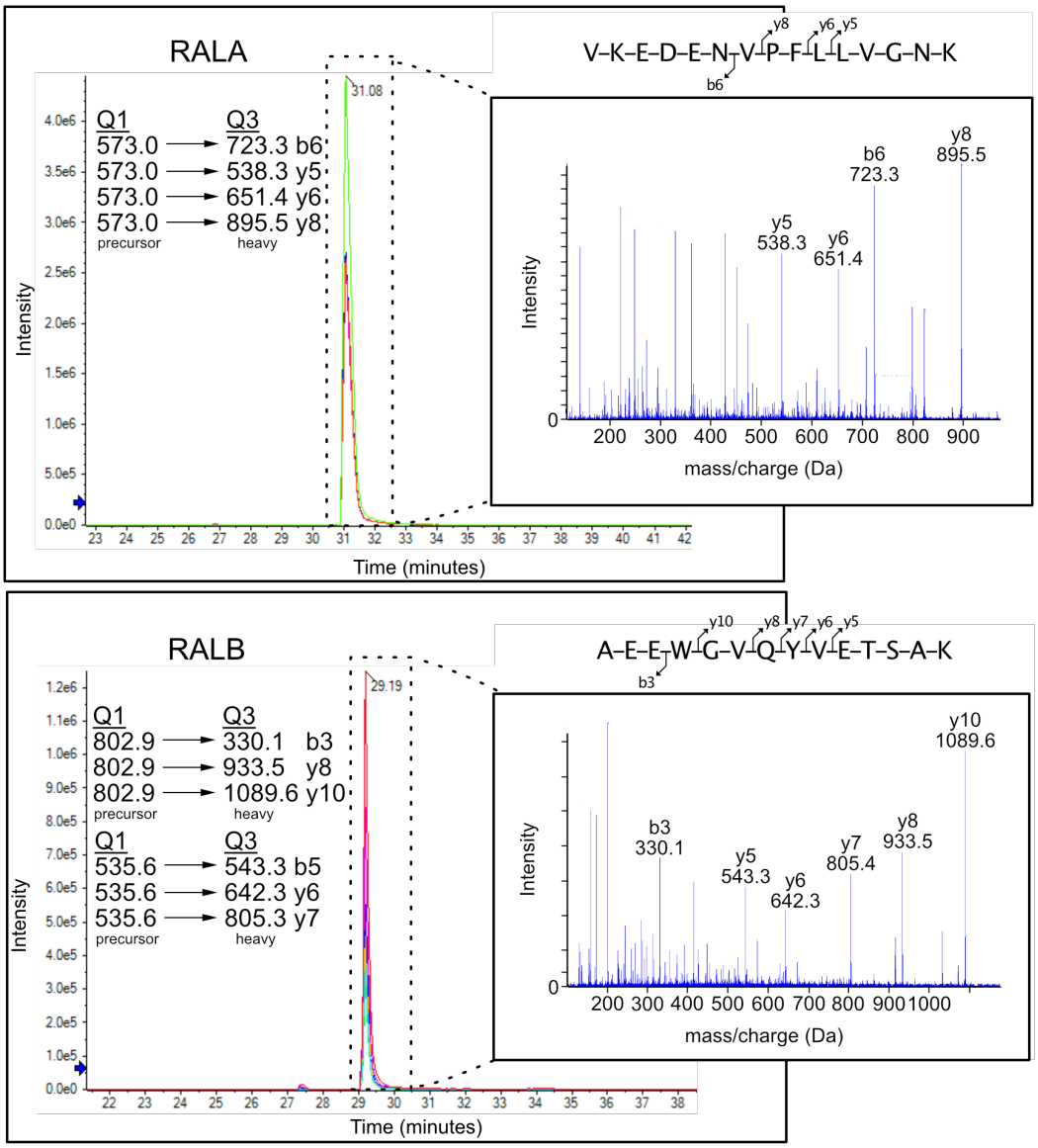
Although detection of all peptides under these conditions seemed promising, identification of the same peptides in subsequent runs in which a range of lower concentrations of heavy recombinant RALA/B (0.5 – 5 ng), with and without lysate, proved to be more challenging (Figure 6.8A). The RALA specific peptide EDENVPFLLVGNK, was again detected with a miscleavage at both the light and heavy levels and instead was detected as VKEDENVPFLLVGNK. Whilst detection of the heavy version of the other RALA specific peptide ADQWNVNYVETSAK was clear, when looking at the light (endogenous) levels, this peptide seemed to increase in a linear manner similar to that of the heavy (data not shown). Combined with the unreliable detection of one RALA specific peptide but the consistent detection of the miscleaved peptide in all samples and subsequent repeat runs, VKEDENVPFLLVGNK was deemed the preferred peptide for future quantification analysis.

Again, identification of the RALB specific peptide IPLLVVGNK was difficult and in fact was undetectable at all lower concentrations as either the light or heavy version. The other RALB specific peptide AEEWGVQYVETSAK was also difficult to identify due to interfering peaks but was the more reliable of the two. Both PAN-RAL candidates were clearly visible in the 100 ng samples but the linearity of each peptide was variable at lower concentrations, in and out of cell lysate. The lack of an identifiable peak for both the light and heavy versions of VIMGSGGVGK and SALTLQFMYDEFVEDYEPTK, in different experimental runs, suggests the peptides may be unstable or that they are being oxidised due to the presence of a methionine in both. Since other runs displayed these heavy peptides as increasing in a linear manner (Figure 6.8A), both peptides were selected to be included in the final calculations and compared to RALA + RALB peptides to provide a level of confidence in peptide detection. Overall, whilst identification of the peptides at these lower concentrations proved to be problematic, the following peptides: VKEDENVPFLLVGNK (RALA), AEEWGVQYVETSAK (RALB), SALTLQFMYDEFVEDYEPTK and VIMGSGGVGK (both PAN RAL) were selected for use in quantifying RAL abundance in my final cell samples.



		PRECURSOR ION (m/z)		PRODUCT ION (m/z)		
PEPTIDE	CHARGE	K0 R0	K8 R10	K0 R0	K8 R10	ION
SALTLQFMYDEFVEDYEPTK	+3	809.4	812.1	345.1	353.2	y3
				752.2	760.3	y6
				881.3	889.4	y7
				980.4	988.4	y8
VIMVGSGGVGK	+2	502.3	506.3	213.1	213.1	b2
				561.3	569.3	y7
				660.4	668.3	y8

Figure 6.6 MS/MS transition spectra for PAN RAL peptides.
100 ng of His-tagged RAL isoforms were subjected to LC-MS/MS analysis for detection of candidate peptides. Both PAN RAL peptides were detected for RALA and RALB. Precursor (Q1) and fragment (Q3) ions are indicated for both the heavy and light peptides. Transitions are chosen based on signal-to-noise and sensitivity to allow robust discrimination versus other peptides with similar mass/charge in the cell lysate. Generally, higher mass range transitions are preferred due to the higher likelihood that they will be selective for the parent peptide.



		PRECURSOR ION (m/z)		PRODUCT ION (m/z)		
PEPTIDE	CHARGE	K0 R0	K8 R10	K0 R0	K8 R10	ION
VKEDENVPFLLVGNK	+3	567.6	573.0	715.3	723.3	b6
				530.3	538.3	y5
				643.4	651.4	y6
				887.5	895.5	y8
AEEWGVQYVETSAK	+2	798.8	802.9	330.1	330.1	b3
				925.5	933.5	y8
				1081.6	1089.6	y10
	+3	532.9	535.6	535.3	543.3	y5
				634.3	642.3	y6
				797.4	805.3	y7

Figure 6.7 MS/MS transition spectra for RALA and RALB peptides.
100 ng of His-tagged RAL isoforms were subjected to LC-MS/MS analysis for detection of candidate peptides. All isoform specific peptides were detected at this concentration with VKEDENVPFLLVGNK (RALA) and AEEWGVQYVETSAK (RALB) selected for future analysis. RALB peptide AEEWGVQYVETSAK was presented with a 2+ and 3+ charge. The 2+ ions were used for data analysis as they displayed less interference.

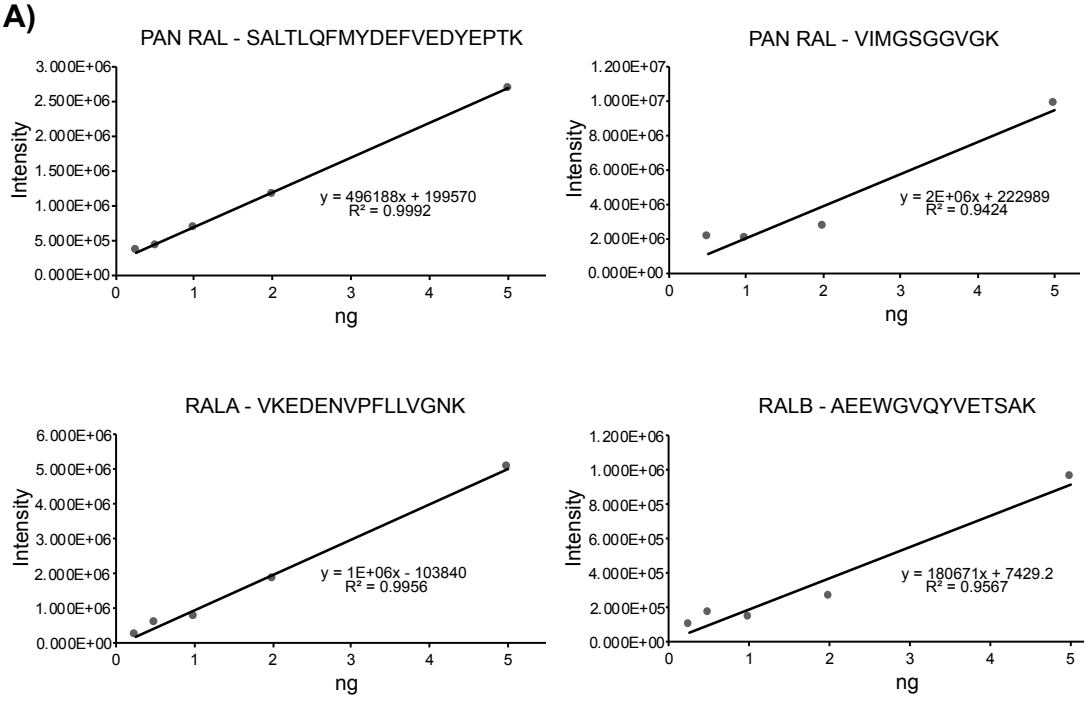
6.2.3 Calibration Curves and determining spike-in concentration

During the selection of the most appropriate peptides for use in quantification, adding a range of standard concentrations (0.5 – 5 ng) into 20 µg cell lysate, allowed rough determination of how much His-RALA/B to spike into the final samples. Using the calibration curve generated from the heavy standards (Figure 6.8A) endogenous peptide readouts from the cell lysate could be compared. Across the five samples analysed in the first run, the average RALA concentration was less than 1 ng, RALB was 3 ng. However, since the PAN RAL (based on SALTLQFMYDEFVEDYEPTK) readout was 1 ng, this concentration was selected for both RALA and RALB for future samples in case the RALB levels in this cell type were higher than average. Following the processing of three batches (21 samples each), it was decided that since the average readout for both the endogenous RALA and RALB in the colorectal cells lines was 2 ng each, future batches would have a spike-in concentration that matched this increase.

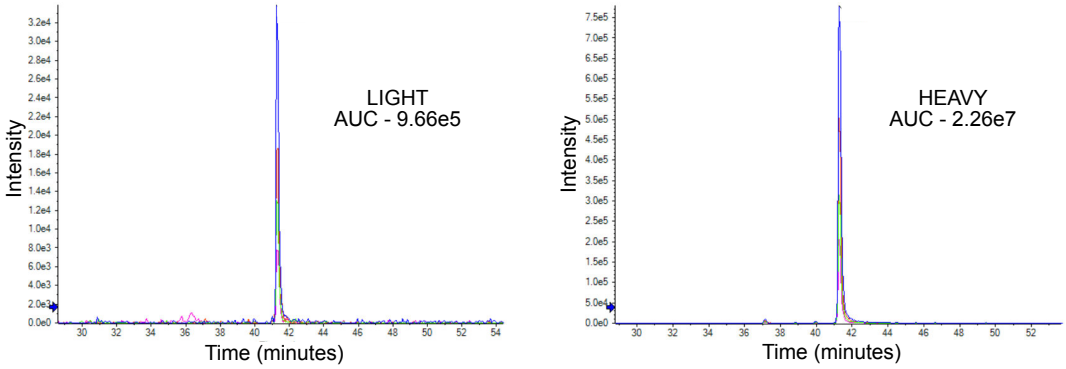
6.2.4 Recombinant protein labelling efficiency

In addition to selecting appropriate peptides, the protein standards were also tested to confirm efficiency of incorporated heavy amino acids (arginine and lysine). Following SDS-PAGE and in-gel digestion, analysis was performed on the mass spectrometer to determine the presence of “light” peptides that would indicate incomplete incorporation (Figure 6.8B). Analysis using LC-MS revealed the percentage of heavy labelling to be 95.7% and 96.3% for the PAN RAL peptides for His-RALA, 95% and 99.2% for the PAN RAL peptides from His-RALB and 99.7 and 99.3% for RALA and RALB specific peptides respectively (Figure 6.8C). Since the percentage of unlabelled standard is low, it is not necessary for this to be accounted for in the final copy number calculations.

A)



B)



C)

			AUC		
isoform	peptide	sequence	light	heavy	% heavy a.a incorporation
RALA	PAN RAL	SALTLQFMYDEFVEDYEPTK	9.66E+05	2.26E+07	95.7
	PAN RAL	VIMVGSGGVGK	3.97E+07	1.06E+09	96.3
	RALA	VKEDENVPFLLVGNK	1.56E+06	5.68E+08	99.7
RALB	PAN RAL	SALTLQFMYDEFVEDYEPTK	4.59E+05	9.19E+06	95.0
	PAN RAL	VIMVGSGGVGK	1.09E+07	1.34E+09	99.2
	RALB	AEEWGVQYVETSAK	4.24E+05	5.79E+07	99.3

Figure 6.8 Protein labelling efficiency and linearity between RAL peptide abundance and MS response.
A) 20 µg HEK293T lysate was mixed with isotope-labelled RALA and RALB at concentration ranging from 0.5, 1, 2, and 5 ng each before being subjected to in-gel digestion. **B)** Representative analysis graph of light and heavy versions of RAL peptides to determine the extent of isotope labelling of full-length PSAQ RAL standards. Low light AUC value in comparison to heavy AUC, indicates high levels of heavy amino acid incorporation. **C)** AUC values for all selected peptides and their percentage heavy amino acid incorporation.

6.2.5 Quantification of endogenous RAL levels in colorectal cell lines

The measurement of cellular RAL abundance covered a range of colorectal cancer cell lines (Table 6.1), including an isogenic panel of SW48 cell lines harbouring a range of RAS mutations. Unlike the other colorectal cell lines in this analysis, these isogenic cell lines were generated through targeted homologous repair in which the endogenous WT RAS allele was exchanged for a mutant RAS sequence. The RAS mutation status of each cell line is also included in Table 6.1 and represents the KRAS mutations present in over 75% of colon tumours (G12D, G12V and G13D), as well as other less prevalent KRAS mutations (G12A, G12C, G12R and G12S) (Prior, Lewis and Mattos, 2012; Vaughn *et al.*, 2011). The SW48 panel of cell lines also included HRAS G12V and NRAS G12V mutations. All mutant cell lines except SW620 were heterozygous for their respective RAS mutations (COSMIC database).

Table 6.1 Panel of colorectal cell lines for PSAQ.

Cell type	RAS mutation status
CaCo2	WT
DLD1	KRAS G13D
HCT116	KRAS G13D
LoVo	KRAS G13D
SW48	WT
SW48	HRAS G12V
SW48	KRAS G12A
SW48	KRAS G12C
SW48	KRAS G12D
SW48	KRAS G12R
SW48	KRAS G12V
SW48	KRAS G13D
SW403	KRAS G12V
SW620	KRAS G12V

Since the cells from each sample were counted and the cell lysate volume measured, it was possible to calculate the number of RAL molecules per cell. The total number of RAL molecules was calculated from the PAN RAL peptide VIMGSGGVGK and varied across the different cell lines (Figure 6.9A). The SW403 cell line had the most RAL out of all the colorectal cell lines tested, with 451,792 ± 114,110 RAL molecules. Comparison with other cell lines containing an active KRAS G12V mutation (SW48 and SW620) revealed noticeably lower values for total RAL with 256,197 ± 52,041 and 169,051 ± 22,328 respectively. This suggests that the difference in RAL expression is not necessarily impacted by KRAS mutation. This is supported by the fact that the second highest RAL total was see in the KRAS WT, CaCo2 cell line with 422,637 ± 99,117 and 349,497 ± 36,899 copies calculated in KRAS G13D HCT116 cells. Although it should be pointed out that the variability between CaCo2 repeats was highly variable, which could be due to the difficulties in handling and counting of this cell line.

That being said, whilst the SW48 WT, SW48 HRAS G12V and SW48 KRAS G13D all display similar levels of total RAL, there was an observable decrease in total RAL in SW48 cell lines harbouring a mutation at codon 12 (Figure 6.9A). Adding to this observation is the fact that the cell line with the lowest number of total RAL molecules, SW620, is also the only cell line within the panel tested to harbour a homozygous KRAS G12V mutation (Figure 6.9A). This decrease in RAL was not associated with a decrease in a particular RAL isoform (Figure 6.9B). This suggests that a KRAS specific codon 12 mutation may impact both RAL isoforms and RAL expression in general.

The overall RAL isoform percentage varies between cell line and RAS mutation (Figure 6.9B). This again implies that RAS is not the main cause of RAL expression level differences. However, a relationship between total RAL expression and RAL isoform % expression was visible (Figure 6.10A). On the whole, those cell lines with higher numbers of total RAL molecules per cell had more RALB expression. Whereas lower total RAL expression correlated with equal RAL expression or in the case of SW620 and LoVo, RALA expression was dominant (Figure 6.10A).

Since there were difficulties in identifying certain peaks during sample analysis, the sum of RALA and RALB molecules was also calculated for

comparison against the PAN RAL values (Figure 6.10B). On the whole these two values were similar, providing some confidence in the reliability of all four peptides in the quantification of RAL. Again, CaCo2 had noticeable variability between PAN RAL and RALA+RALB values (Figure 6.10B). This is likely a result of the cell line handling issues mentioned earlier, as well as difficulties in peptide peak detection during data analysis.

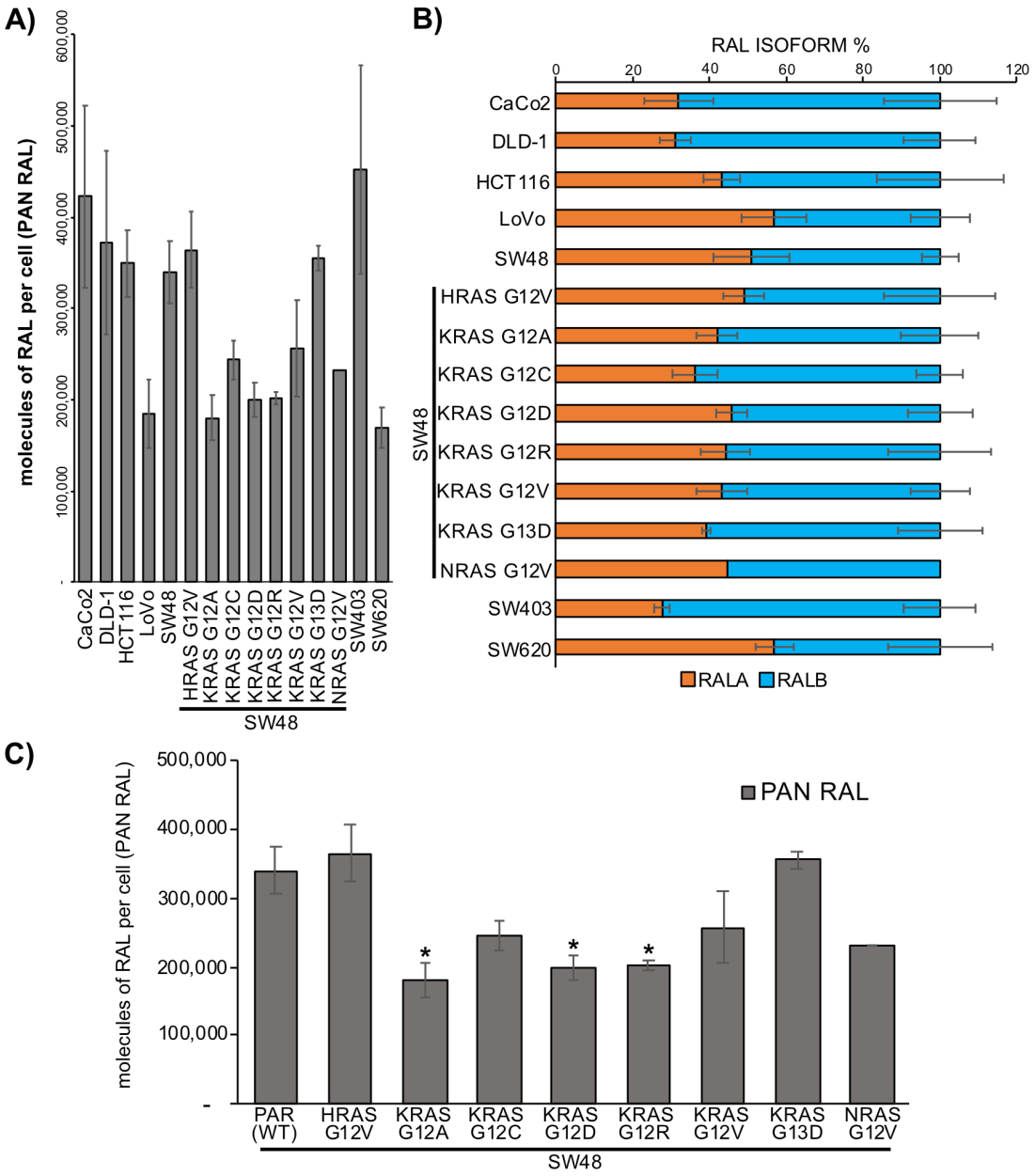


Figure 6.9 Cellular RAL abundance in colorectal cell lines.
A) Total number of RAL molecules per cell (PAN RAL) in different colorectal cell lines including a panel of SW48 isogenic cell lines with various RAS mutations. **B)** RAL isoform percentage within each cell line. **C)** Total number of RAL molecules per cell (PAN RAL) across a panel of SW48 isogenic cell lines with knock-in RAS mutations. A decrease in RAL expression was seen in cell lines harbouring a KRAS mutation at codon 12 vs SW48 WT. KRAS G12A, G12D and G12R were significantly decreased ($p < 0.05$). All bars represent mean ± SE of 3 biological repeats, except SW48 NRAS G12V that is n=1.

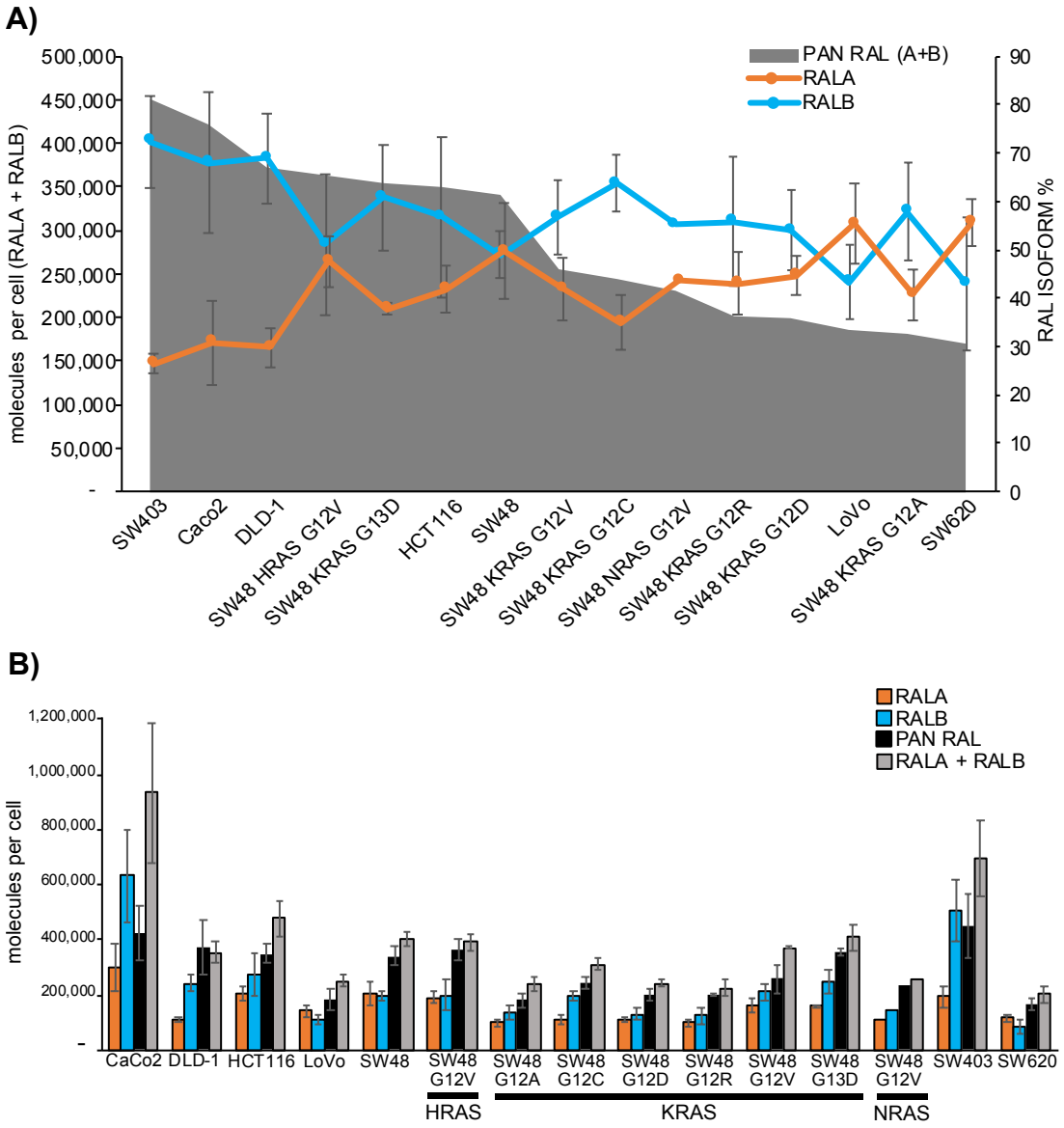


Figure 6.10 RAL isoform abundance in colorectal cancer cell lines.
A) Percentage of RAL isoforms present in colorectal cancer cell lines, in relation to total number of molecules per cell (RALA + RALB), ordered highest to lowest RAL abundance. All data represent mean ± SE of 3 biological repeats, except SW48 NRAS G12V that is n=1.
B) number of molecules of RALA and RALB across a panel of colorectal cell lines including SW48 isogenic knock-in mutants. Total RAL abundance per cell is also included, along with values for RALA + RALB for comparison.

6.3 Discussion

The PSAQ method is based on the inclusion of full-length recombinant proteins that can be cleaved to produce proteotypic peptides that uniquely identify the protein(s) of interest. Candidate peptides are selected based on several factors including sequence length, likely charge states, known posttranslational modifications (PTM) and the presence of modifiable residues. Since many proteomic studies are based on tryptic peptide identification, in this instance trypsin was also utilised to proteolyse the standards and generate both PAN RAL and isoform specific RAL peptides. Whilst this was sufficient in generating peptides that were then successfully used to calculate RAL abundance, there were factors associated with each peptide, especially the PAN RAL peptides, that hindered the overall process.

Any PTMs on RAL must be taken into account when selecting appropriate peptides for quantification. Both RAL isoforms undergo geranylgeranylation in order for proper membrane localisation to take place (Falsetti *et al.*, 2007). In addition to these lipid modifications, both RAL isoforms can undergo phosphorylation at differing serine residues to allow for subcellular relocation and function (Wu *et al.*, 2005; Wang *et al.*, 2010; Sablina *et al.*, 2007; Kashatus *et al.*, 2011; Martin *et al.*, 2012). All of these modifications occur in the hypervariable region of RAL and therefore any peptides covering this region were discounted. RAL isoforms can also be modified on residues outside of this region and have both been found to be monoubiquitinated in a nondegradative manner (Neyraud *et al.*, 2012). Several lysine (K) residues are viable sites for ubiquitination. In particular ubiquitination of K47 in RALB determines effector binding and subsequent functionality (Simicek *et al.*, 2013). This lysine residue is situated at the end of the PAN RAL peptide SALTLQFMYDEFVEDYEPTK. The extent of RALB ubiquitination at K47 is unclear, although arginine substitution markedly reduced RALB ubiquitination levels, so addition of the second PAN RAL peptide VIMGSGGVGK was useful in validating PAN RAL levels. Ubiquitination of RALB has also been detected at K129 and K160 residues at the end of both RALB specific peptides IPLLVVGNK and AEEWGVQYVETSAK (Simicek *et al.*, 2013). Unfortunately, these were the only two candidates produced following trypsin digestion that

were deemed suitable. Therefore, other proteases may be able to provide alternative peptides that can complement the tryptic digest and be utilised for future sample analysis if necessary.

Nearly all the peptides generated included additional residues that can be chemically modified, which presents a problem for SRM-based strategies. Both PAN RAL peptides included a methionine (M) residue and one also contained a glutamine (Q). Both RALA peptides included an asparagine (N) residue and one RALA and one RALB-specific peptide also included a glutamine. Methionine can undergo modification in the form of oxidation, generating methionine sulfoxide or sulfone, resulting in a mass increase of 15.99 or 31.99 respectively (Ghesquière and Gevaert, 2014). Glutamine and asparagine residues can undergo deamidation and be converted into glutamate and isoaspartate/aspartate respectively (Geiger and Clarke, 1987; Lange *et al.*, 2008; Piszkiewicz, Landon and Smith, 1970), causing an increase in mass of almost 1 Da. Changes in mass such as this are problematic, since modified peptides will be ejected if not specifically selected for in SRM. Chemical modifications will also act to spread the signal of a peptide, reducing the sensitivity of the assay, which in this instance is already utilising low concentrations of protein standards.

The use of the PSAQ strategy ought to compensate for any of these acquired modifications, since the isotope-labelled standard is added to the sample at the earliest possible point in the workflow and therefore should be equally affected as the endogenous proteins. However, other common proteases such as LysC, elastase and GluC might be worth consideration for their ability to generate proteotypic peptide that overcome these problems. In fact, a simulated digestion with GluC, which cleaves at the C-terminal end of glutamic acid (E) residues, revealed two alternative PAN RAL peptides (DYEPTKADSYRKKVVLDGEE, TSAKTRANVDKVFFDLMRE) and two RALB specific peptides (DKIPLLVVGNKSDLEE, GFLLVFSITE). However, these PAN RAL peptides contain multiple mid-chain highly basic residues that may limit their fragmentation and dilute the response signal. Furthermore, one of the RALB specific peptides includes the problematic tryptic peptide IPLLVVGNK that was completely undetectable under the conditions used in this assay.

Besides further optimisation of peptide generation and selection, other steps could be taken to enhance peptide detection and increase reliability in peptide peak presence. It was highlighted early on in the PSAQ process that contamination from residual Triton X-100 lysis buffer and/or acrylamide within the gel was interfering with data acquisition on the mass spectrometer. This meant that optimisation steps to reduce the volume of gel excised, whilst still having sufficient material for MS analysis was required. In addition to using an alternative NP40 based lysis buffer, lowering the acrylamide percentage, running the gel just far enough to get clear resolution of the protein ladder, as well as using small, narrow lanes, reduced the size of the gel piece excised eliminated the problems initially seen when processing samples. As such only 20 µg of cell lysate with ng volumes of standards were used in the final sample preparations. Using lower concentrations also reduces the visibility of peptide peaks and increases background noise. During data analysis the parameters can be adjusted to display the retention times for lower intensity peaks. Whilst this was necessary to detect the low concentration of standards and endogenous protein levels involved, it should be pointed out that quantifying peaks below the limit of quantification is not advisable. This limitation is defined in various ways, with the most common one being the lowest calibration standard with a percentage coefficient of variation (CV) less than 20. Whilst almost all the intensity peaks were clear enough to provide a certain degree of confidence during peptide identification, upscaling the amount of cell lysate and spike-in standard concentration to those above the normal detection values, in combination with an in-solution digest protocol to eliminate gel contaminants would be desirable.

Regardless of any issues in peptide detection and sample preparation, the presented data represents the most accurate quantification of cellular RAL isoform abundance to date. By calculating RAL protein expression across a variety of relevant human cancer cell lines, insight was gained into the factors that may impact RAL-dependent signalling and oncogenic RAS.

Overall, the total RAL expression differed greatly from line to line, with over 500,000 copies per cell difference calculated between the highest and lowest recorded values. Excluding the CaCo2 cell line, the two cell lines with the

highest RAL abundance were SW403 and HCT116. This implies the presence of highly active RAL signalling in these cell lines, as well as placing a certain degree of dependency on the RALGEF/RAL pathway in carrying out various functions. Indeed, out of a panel of colorectal cell lines including the same ones analysed in this study, HCT116 and SW403 displayed the strongest decrease in cell viability following RAL knockdown (Lee *et al.*, 2019). As well as high total RAL expression, these cell lines reported high RALB levels as well as RALB being the dominant isoform expressed. This link between cell viability, high total RAL and high RALB abundance is consistent with the findings that RALB was found to be more important for cell survival, with transformed cell lines undergoing apoptotic cell death following RALB knockdown (Chien and White, 2003). In addition to HCT116 and SW403, RALB was found to be the dominant isoform across the cell lines tested or was in equal abundance to RALA. Indeed, the RAL activity assay performed in the previous chapter support the notion that a certain level of both RALA and RALB is required in these cells, since activity increases in the remaining RAL were detected following isoform specific knockout, across all cell lines tested. This is in contrast to RAL expression in pancreatic cell lines, in which RALA was found to be the dominant isoform expressed and RALB was in fact undetectable in pancreatic islets (Ljubicic *et al.*, 2009). This highlights the differences in RAL requirements between cell types and needs to be taken into consideration when investigating any role that RAL might play in the dysregulation of cellular processes.

The inclusion of the SW48 isogenic cell lines allowed us to determine the effect, if any, of different RAS mutations on RAL protein abundance. The first and second most abundant KRAS mutation in colorectal cancer are G13D and G12V respectively (Vaughn *et al.*, 2011). As previously mentioned, cell lines SW403 and HCT116 that carry these mutations, displayed high RAL abundance supporting a link between mutant RAS and increased RAL expression levels. Interestingly, SW48 cell lines harbouring these same mutations did not induce significant changes in RAL expression. However, it is important to point out that the SW48 cell line is not dependent on RAS to be viable and therefore may not be a true representation of RAS driven signalling.

In relation to mutant RAS and oncogenic activity, HRAS G12V mutants that activate different RAS effector pathways, have shown that it is the RALGEF pathway that is sufficient for RAS transformation in human cells (Hamad *et al.*, 2002). Additionally, early RAS research shows HRAS G12V mutation have the greatest potency compared to other HRAS mutations (Der, Finkel and Cooper, 1986; Seeburg *et al.*, 1984; Sloan, Newcomb and Pellicer, 1990). Furthermore, similar protein abundance studies currently being conducted in our lab found that the SW48 HRAS G12V cell line had more total RAS than the SW48 parental and other mutation cell lines (Hood *et al*. unpublished). Therefore, it is perhaps surprising that the overall RAL expression in the SW48 HRAS G12V cell line has not deviated from that recorded in the SW48 parental, neither is there one RAL isoform whose expression dominates the other. Again, this lack of change in RAL abundance suggests that mutant RAS does not impact RAL expression and consequently does not contribute to RAS oncogenicity. However, as well as the limitations associated with this cell lines, it is important to remember that this assay is only indictive of protein expression not activity. Therefore, it is beneficial to refer to the activity assays performed in the previous chapter that showed the RAL knockout cell lines with mutant RAS had significant activity compensation compared to SW48 RAS WT. Since mutant RAS expression is associated with oncogenic transformation, any involvement of the RALGEF/RAL pathway may be due to changes in activity, not expression. Moreover, the transformation potential of different RAS effectors reportedly varies in rodents depending on cell type (Cowley *et al.*, 1994; Mansour *et al.*, 1994; Oldham *et al.*, 1996). It is possible that transformation of human cells seen using RALGEF specific RAS effector mutants may be restricted to specific cell types and determining RAL abundance and activity levels in cancer cell lines with alternative tissue origins is desirable. Again, focusing on the SW48 isogenic cell lines, with the exception of HRAS G12V and KRAS G12V, those with a mutation at codon 12 displayed reduced RAL expression compared to SW48 parental. This implies that G12 RAS mutations may actively effect RAL abundance and in fact suppress RAL-driven processes.

In relation to G12V mutations, even though cell lines harbouring a KRAS G12V mutation were found to have some of the highest levels of RAL expression,

SW620 cells that also have this mutation contain the lowest RAL abundance out of any of the cell lines tested. This is despite SW620 being homozygous for mutant KRAS. However, since RAS mutations induce cellular stress (Grabocka and Bar-Sagi, 2016), a homozygous KRAS G12V mutation could produce such a prolific oncogenic signal that the cells actively limit its cellular abundance and consequently RAL expression.

In general, this analysis of RAL isoform abundance suggests that regardless of RAS status, high RAL abundance contributes to a strong RAL signalling response and varying RAL expression levels is determined by the overall expression profile of the cell. Nevertheless, the data collected here indicates that mutant RAS can influence RAL, although it appears that this is not solely confined to changes in expression but may present as a change in RAL activity instead. It seems that for the most part, colorectal cancer cell lines are reliant on both RAL isoforms, with a small bias towards RALB expression. This is reflected in the activity assays. Indeed, when combined with the RAL activity data collected, it is likely that abundance and activity work together to carry out overall RAL function in colorectal cancer.

Chapter 7: Outlook and future work

The small GTPase RAL is known to have a role in many different cellular functions (Feig, 2003; Gentry *et al.*, 2014). That being said, the majority of our understanding of RAL signalling has been inferred from knockdown/ectopic expression studies, often in mouse or human cell lines that are not RAS addicted or prone to RAS-dependent malignancies. Furthermore, whilst several distinct functions can be associated with each RAL isoforms, isoform specific RAL biology remains a relatively unexplored area and outside of its immediate effectors, the RAL pathway is still poorly understood in comparison to other RAS effector pathways. Therefore, this study focused on investigating isoform specific RAL signalling at the endogenous level in an appropriate human cell model.

In the initial stages of this project, I successfully generated isoform-specific RAL knockout clones across several colorectal cancer cell lines using the CRISPR/Cas9 system, as well as optimised the tools necessary to establish gain-of-function knock-in cell lines.

The discovery of the CRISPR/Cas9 system has revolutionised gene editing across many areas of biomedical research including model cell line development. Its increasing popularity and widespread use can be explained through its simplicity and flexibility to target nearly any location within the genome. Although there are many advantages to other gene editing tools, such as ZFNs and TALENs, they lack the versatility associated with CRISPR and can often require more complex expertise in molecular biology. However, the CRISPR/Cas9 system suffers from its own set of drawbacks, particularly those focusing on its efficacy that remain to be fully addressed. These factors include on- and off-target cleavage, guide design, DNA repair (NHEJ and HDR) efficiency and delivery of the required components. These last two points in particular were challenges faced in this project when trying to establish knock-in cell lines. As well as working with the low knock-in efficiency associated with HDR, low transfectability of target cell lines hindered this process and steps were taken to optimise this strategy to aid component delivery into cells that are hard to transfect. As a result of these technical

issues and subsequent time constraints, RAL G23V knock-in cells were not generated. In addition to complementing the data obtained from the knockout cell panel, establishing gain-of-function cell lines would enable the investigation of constitutively active RAL at endogenous levels and any part it may play in the role of oncogenic RAS. Nevertheless, the capability of the CRISPR/Cas9 system to precisely edit RAL gene sequences and halt gene expression by generating RAL knockouts has been demonstrated. Studying protein function through CRISPR/Cas9 generated gene knockouts is advantageous over RNAi-based knockdowns, as it is a permanent change that guarantees 100% knockout efficiency due to the selection of a single clone. Currently, these clones are the first to be created for the use of investigating endogenous RAL isoform signalling in a cancer relevant panel of human cell lines.

A combination of methodologies was applied to this cell panel, in order to obtain an insight into cellular RAL isoform signalling under different oncogenic RAS mutations. Determining RAL activity through pulldown of GTP-bound RAL revealed a compensatory increase in activity (significant RALB activity increase in RALA KO clones harbouring mutant RAS) in the remaining RAL isoform in the knockout clones compared to the parental clones. This is in concordance with a study conducted in mice cells, that also showed activity compensation following single RAL isoform deletion (Peschard *et al.*, 2012). This compensatory increase raises the question of whether there is functional redundancy between the two isoforms. Indeed, this same study demonstrated deletion of both RAL GTPases impaired cell proliferation, whilst single gene knockout did not. Hence, isoform activity compensation suggests that the overall level of RAL activity is a crucial factor in RAL signalling response, not just the expression levels of a specific isoform. Going forward it would be beneficial to extend these findings to the SW48 isogenic cell lines to determine the direct impact of RAS mutational status on RAL activity. It would also be beneficial to test these functional implications through proliferation assays in our panel of cell lines to confirm this redundancy in a human model. Moreover, these assays could be applied to knock-in cell lines to help characterise

functional responses associated with constitutively active RAL which mimics signalling dysregulation.

The observation of isoform compensation was subsequently also seen in the cell lines that underwent omic profiling. Whilst the expression of genes and proteins in the KO clones compared to the parental were often different between the two cell lines, changes in expression nearly always trended in the same direction between RAL isoforms. An interesting observation was the lack of responses within the immediate RAL network. Whilst the RPPA assay did not cover many of the interactors involved in the RAL pathway, the RNAseq data and immunoblotting validation failed to highlight any significant changes in known RAL effector proteins. Since the RAL isoforms interact with many of the same set of proteins, it's likely that the compensation in activity carries over into effector interactions and therefore little change in expression of effectors is not surprising.

Nevertheless, this unbiased multi-omic approach has allowed us to study RAL isoform specific signalling in a wider context. Indeed, RNA sequencing and RPPA analyses revealed a subset of RAL responsive proteins that were suggestive of crosstalk between the RALGEF/RAL pathway and the other main pathways downstream of RAS (MAPK and PI3K). Out of the several proteins shortlisted for further investigation, differential expression of DUSP5, EphA2 and WEE1 was confirmed by immunoblotting. However, further validation is needed to characterise these proteins as part of RAL signalling. Of these, only a link between RAL and EphA2 has been recorded in previous studies. It seems likely that RAS signalling is based on a complex network of nodes rather than just the classical linear pathways that are more familiar. Potential crosstalk between pathways also presents the idea of signalling fail safes that ensure RAS-driven functionality, that could be taken advantage of by oncogenic RAS especially following therapy-induced resistance. Therefore, it would be desirable to subject knock-in RAL mutant cells to assays (with and without RAS pathway inhibitors) that would determine invasiveness, proliferation and cell viability, to ascertain the impact of RAL dysregulation.

However, analysis at both the transcriptome and proteome level also revealed how different the expression profiles of cell lines can be, even if they have the same tissue origin. Both cell lines appeared to activate distinct sets of proteins,

suggesting differential signalling occurs downstream of RAS and therefore RAL, depending on the overall profile of the cell. This highlights how important it is to choose the right cell lines when investigating cell signalling proteins such as these. It also shows how regardless of the similarities in mutation status of direct RAL regulators, the overall expression of the proteome is the overriding factor that impacts the signalling response of RAL and how this pathway interacts as part of a larger network. Future studies in this field need to be aware of these concepts and not extrapolate from an observation in a single cell line. Indeed, large-scale studies looking at endogenous mutant KRAS signalling dependencies showed significant heterogeneity in effector requirements between cell lines with each line requiring a unique combination of effector interactions (Yuan *et al.*, 2018). This is also an important consideration when inferring a role for RAL in the oncogenicity of RAS. Although RAL can be implicated in a variety of RAS-driven functions such as cell proliferation and survival, the contribution and dependency of RAL signalling is likely to differ on a case by case basis.

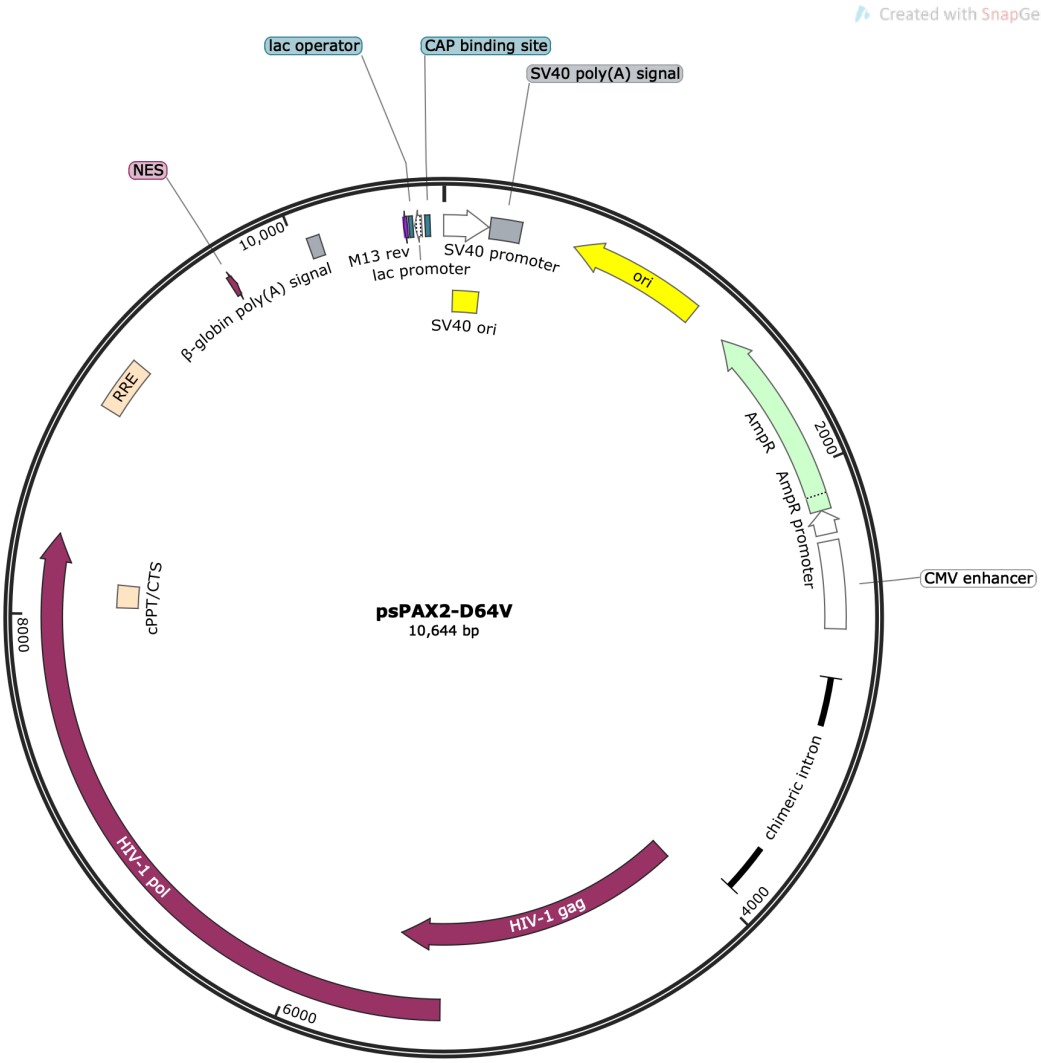
Finally, the SRM-based quantification method in combination with PSAQ standards has enabled the most accurate measurement of cellular RAL isoforms to date. Following adaptation and optimisation of this technique for use against RAL, it reliably detected both RALA and RALB at endogenous levels, as well as peptides shared between both isoforms to verify the isoform-specific quantification. Moreover, the PSAQ method accounted for variations in the recovery of peptides due to gel excision difference and incomplete proteolysis. Importantly, this technique is easily implementable in most biological laboratories due to the familiarity with SDS-PAGE fractionation and lack of specialist knowledge required for the in-gel digestion steps. As a result, this technique would easily enable the quantification of cellular RAL abundance across any cell line and even tissue samples.

In this instance, abundance was calculated in a variety of different colorectal cancer cell lines, as well as in an isogenic panel of SW48 cells. It seems that for the most part, colorectal cancer cell lines require sufficient levels of both RAL isoforms for proper signalling, with a small bias towards RALB expression. This again supports the idea of high levels of functional compensation between the two isoforms in colorectal cancer, with any

differences in RAL isoform function resulting from changes in location and posttranslational modifications. Finally, looking at the abundance dataset in conjunction with the isoform activity analysis, it seems that the mutation status of RAS may directly influence RAL, affecting a combination of abundance and activity levels.

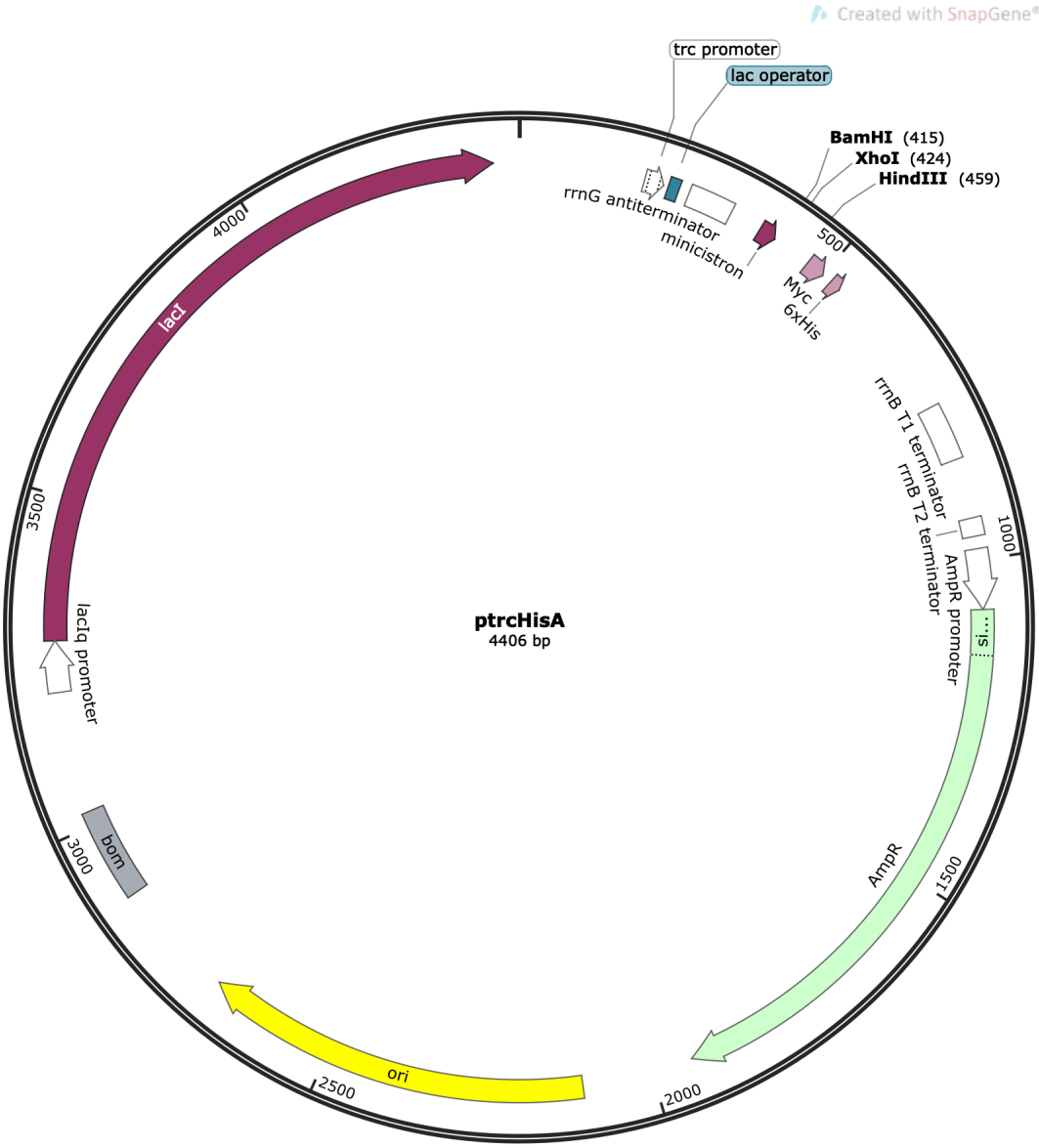
Together, the data presented in this thesis allows us to build knowledge of the specific set of requirements needed in both normal and pathophysiological RAL signalling. Knowing how these different factors combine together, helps to place the extent of RAL's biological effect as part of a larger signalling network.

Supplementary Material

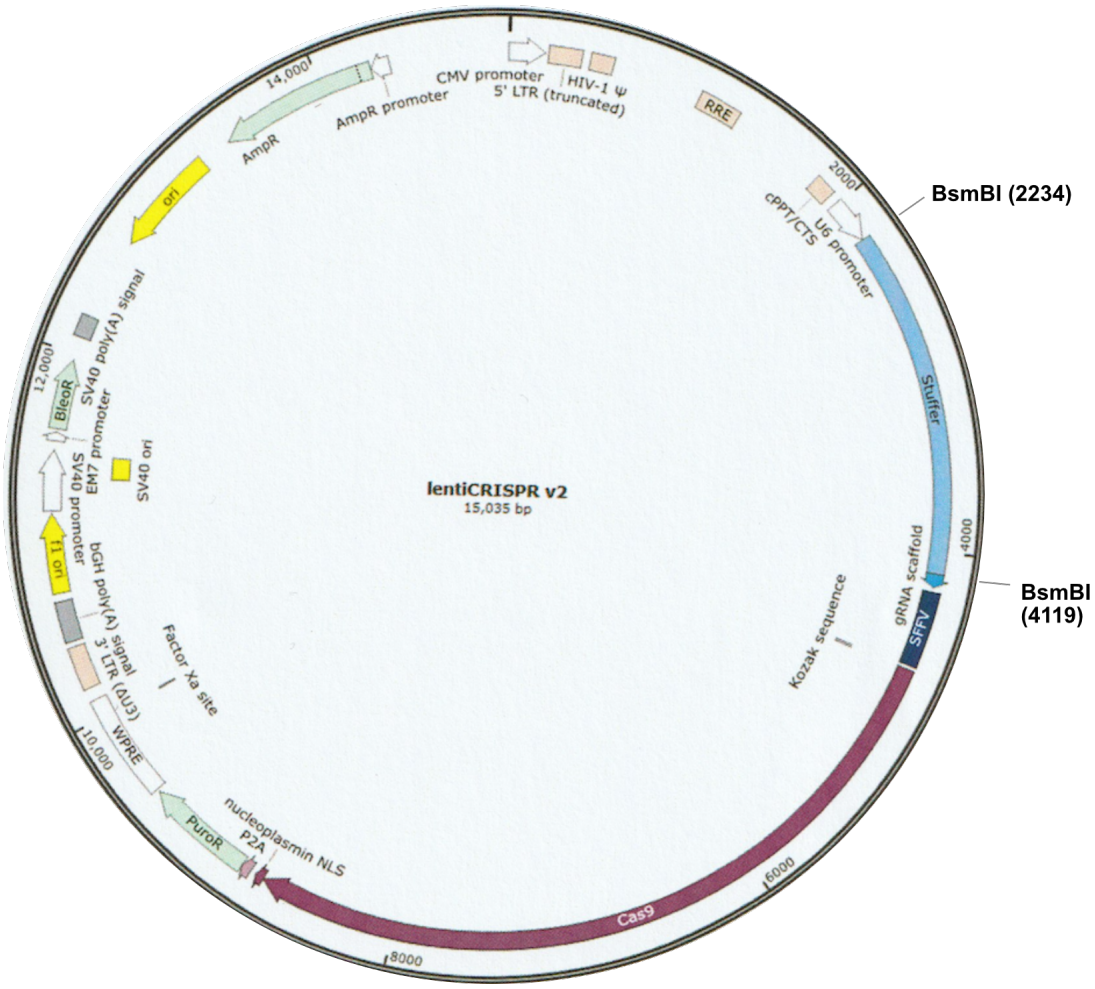


Supplementary Figure 7.1 psPAX2-D64V plasmid map.
The psPAX2-D64V plasmid was obtained from Addgene (#63586) and is a packaging plasmid used for generating integrase deficient lentiviral vectors due to a point mutation in the integrase gene. This stops any fluorescent markers from the transfer plasmid, such as the EGFP present in the pLeGO plasmid, from inserting into the target genome which would prevent any future assays involving fluorescent analysis.

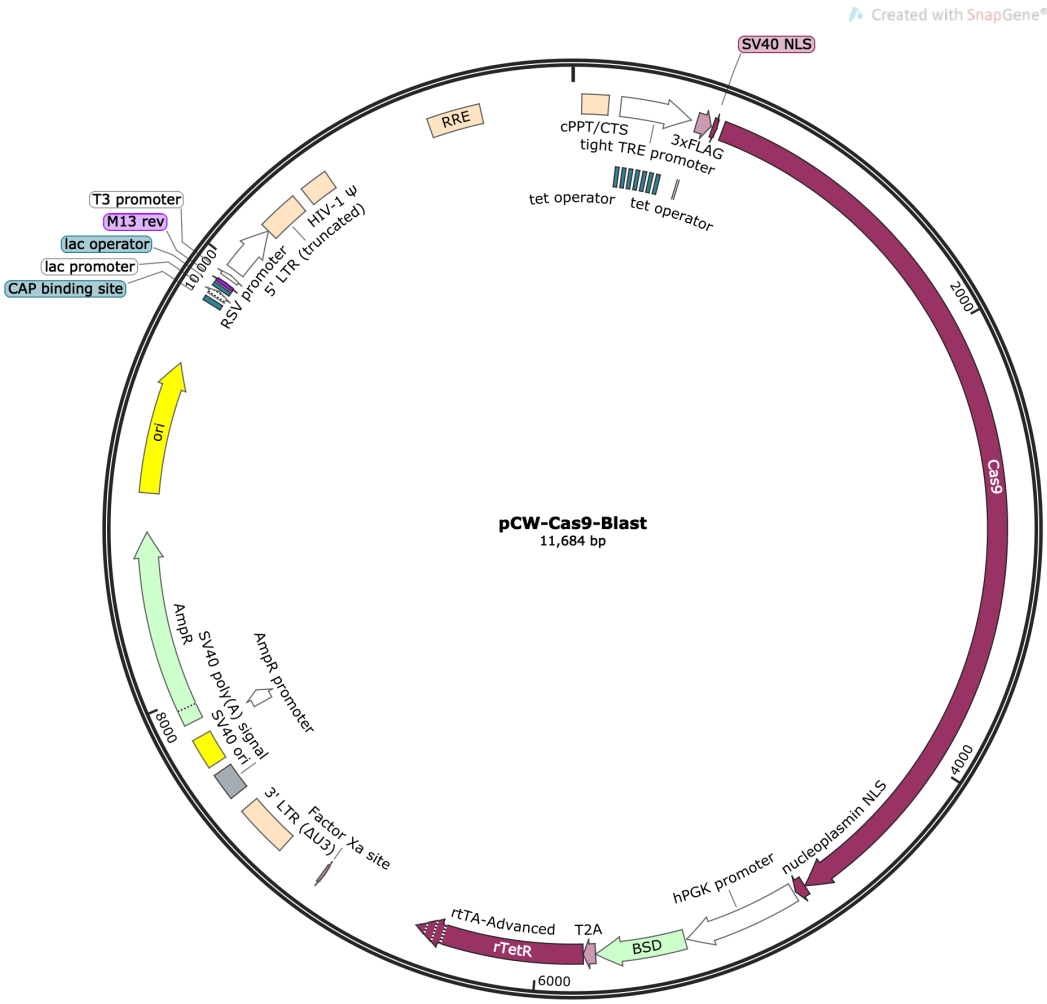
Supplementary Figure 7.2 pBABE-puro plasmid map.
All the WT and mutant RAL constructs were derived from the pBABE-puro plasmid (Addgene #1764). All RAL sequences were inserted between BamHI and SalI (RALA) and BamHI and XhoI (RALB).



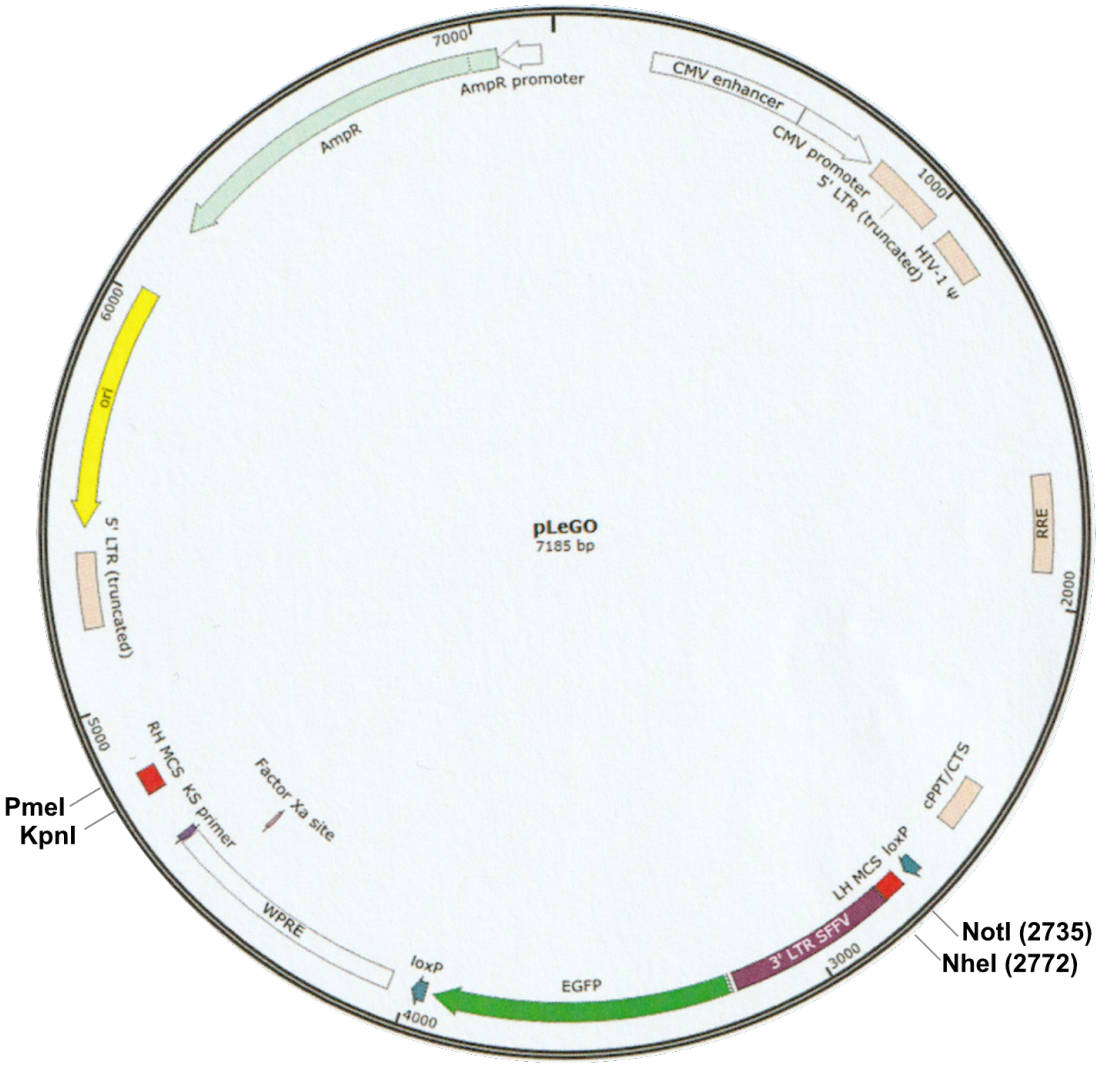
Supplementary Figure 7.3 ptrcHisA plasmid map.
The ptrcHisA plasmid was used to generate His-tagged WT RALA and WT RALB recombinant proteins for PSAQ. Importantly, this plasmid contains the sequence for the LacI gene that is required for successful IPTG induction and subsequent production of T7 RNA polymerase for gene transcription. This gene is absent in the AT713 bacteria used to generate heavy-labelled proteins.



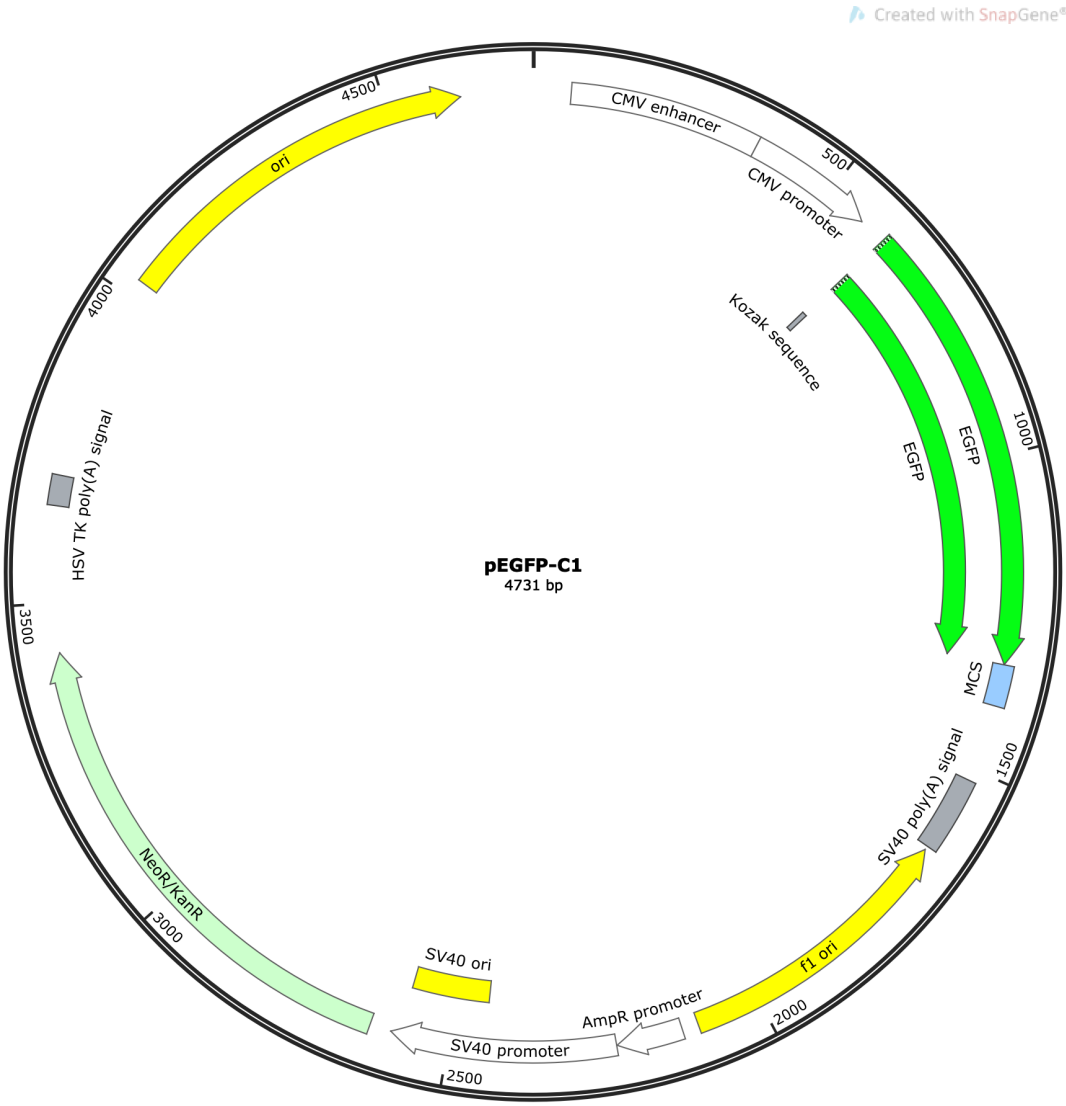
Supplementary Figure 7.4 LentiCRISPR V2 plasmid map.
This plasmid is a lentiviral SpCas9 expressing plasmid purchased from Addgene (#52961). The EF-1α core promoter in the original plasmid was replaced with a SFFV promoter. To allow cloning of gRNAs, the plasmid was digested with BsmBI to remove the stuffer region before ligation of guide oligos into the plasmid.



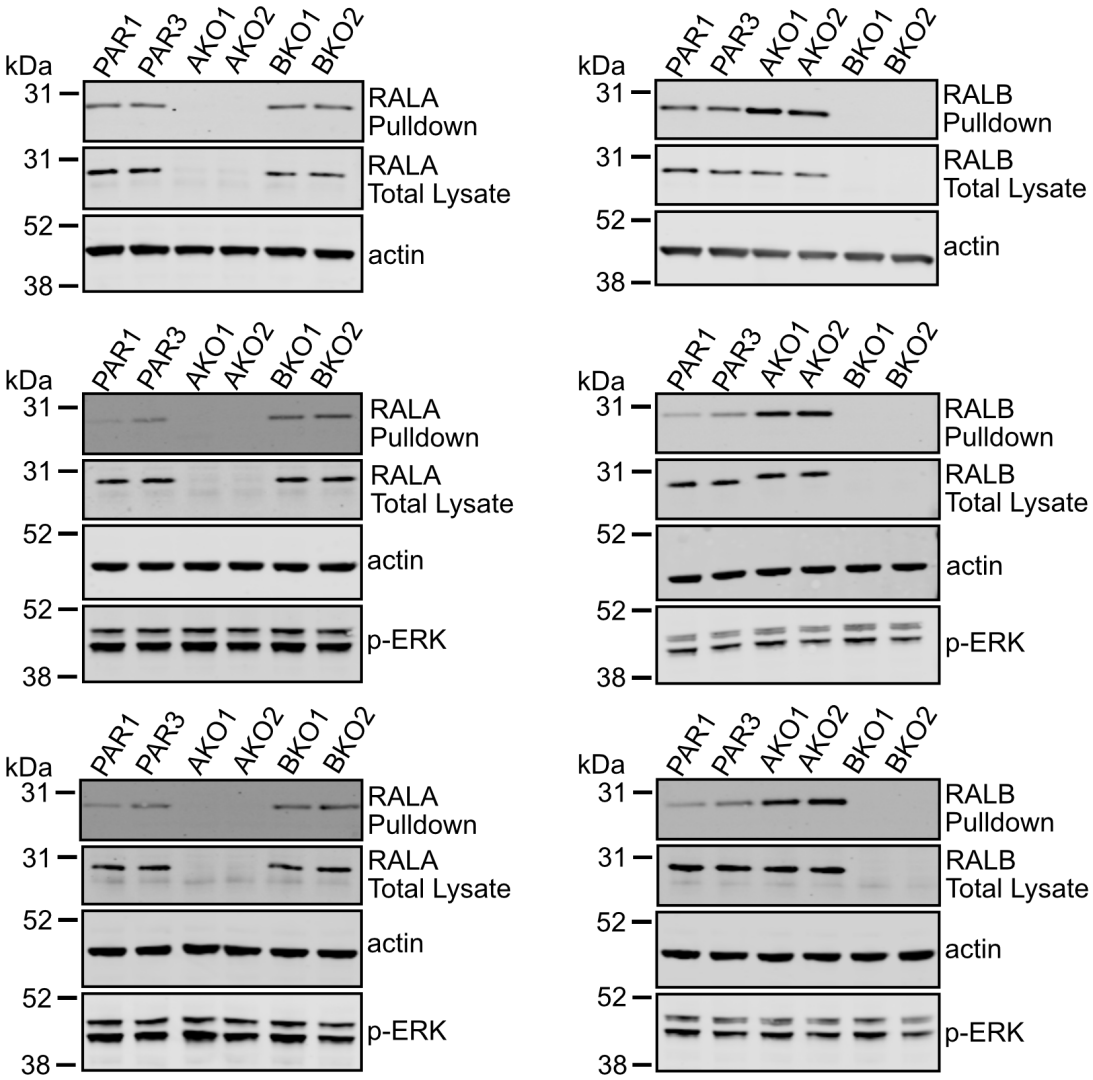
Supplementary Figure 7.5 pCW-Cas9 plasmid map.
This plasmid expresses a doxycycline inducible lentiviral SpCas9 with a blasticidin antibiotic selection marker. The Cas9 enzyme is under the control of a Tet-ON promoter. Purchased from Addgene (#83481).



Supplementary Figure 7.6 pLeGO plasmid map.
The pLeGO lentiviral plasmid was initially purchased from Addgene (#27341) and expresses EGFP under the control of an SFFV promoter. For CRISPR use, this plasmid was modified to contain two multiple cloning sites (MCS, RH and LH depicted in red) that sit either side of the EGFP selection marker. The additional cloning sites allowed cloning of gRNA#4 into the RH MCS using PmeI and KpnI, as well as the HDR template into the LH MCS using NotI and NheI.

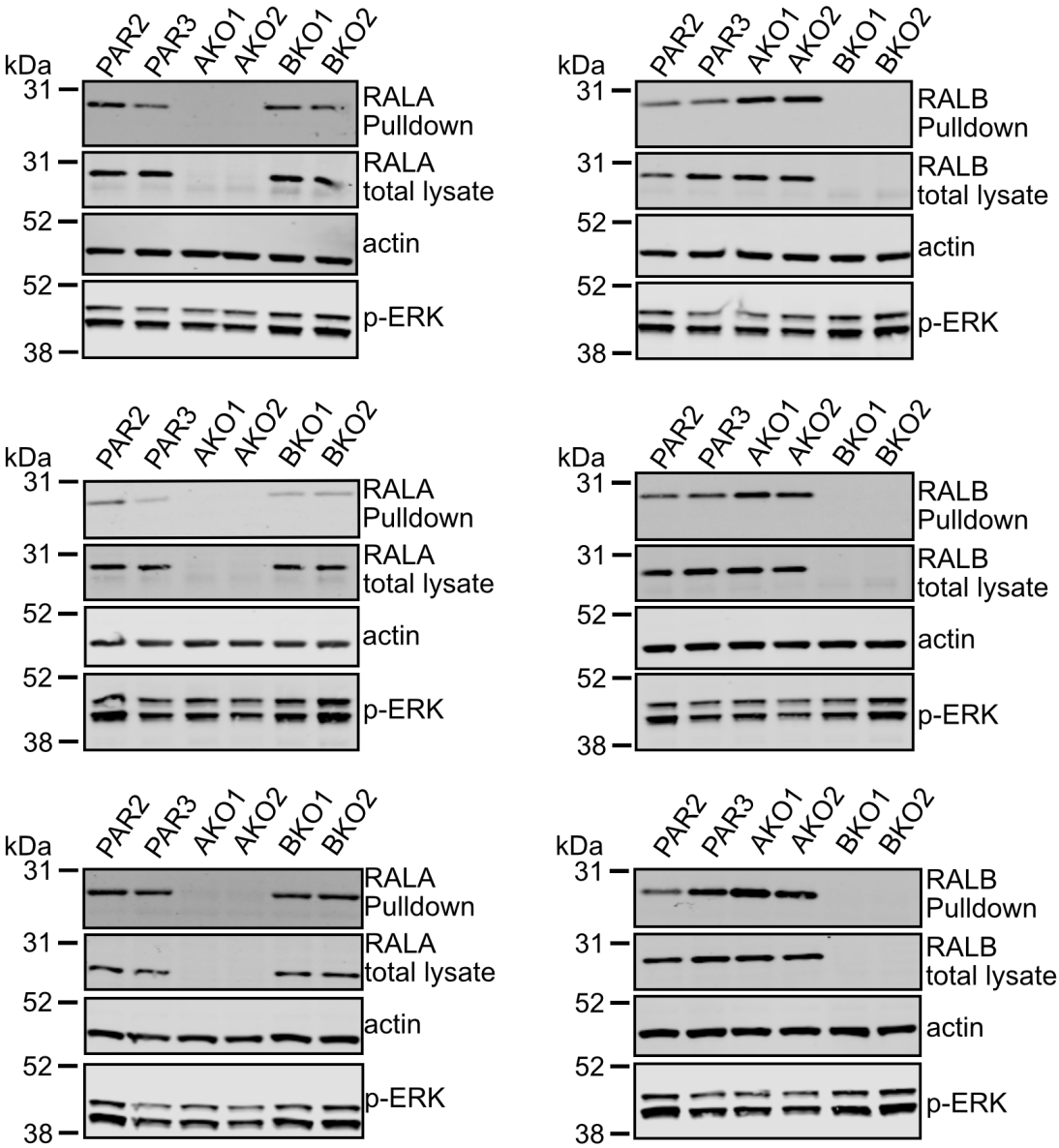


Supplementary Figure 7.7 pEGFP-C1 plasmid map.
The pEGFP-C1 plasmid expresses an EGFP fluorescent marker under the control of a CMV promoter. This plasmid was used to assess transfection efficiency and plasmid uptake.



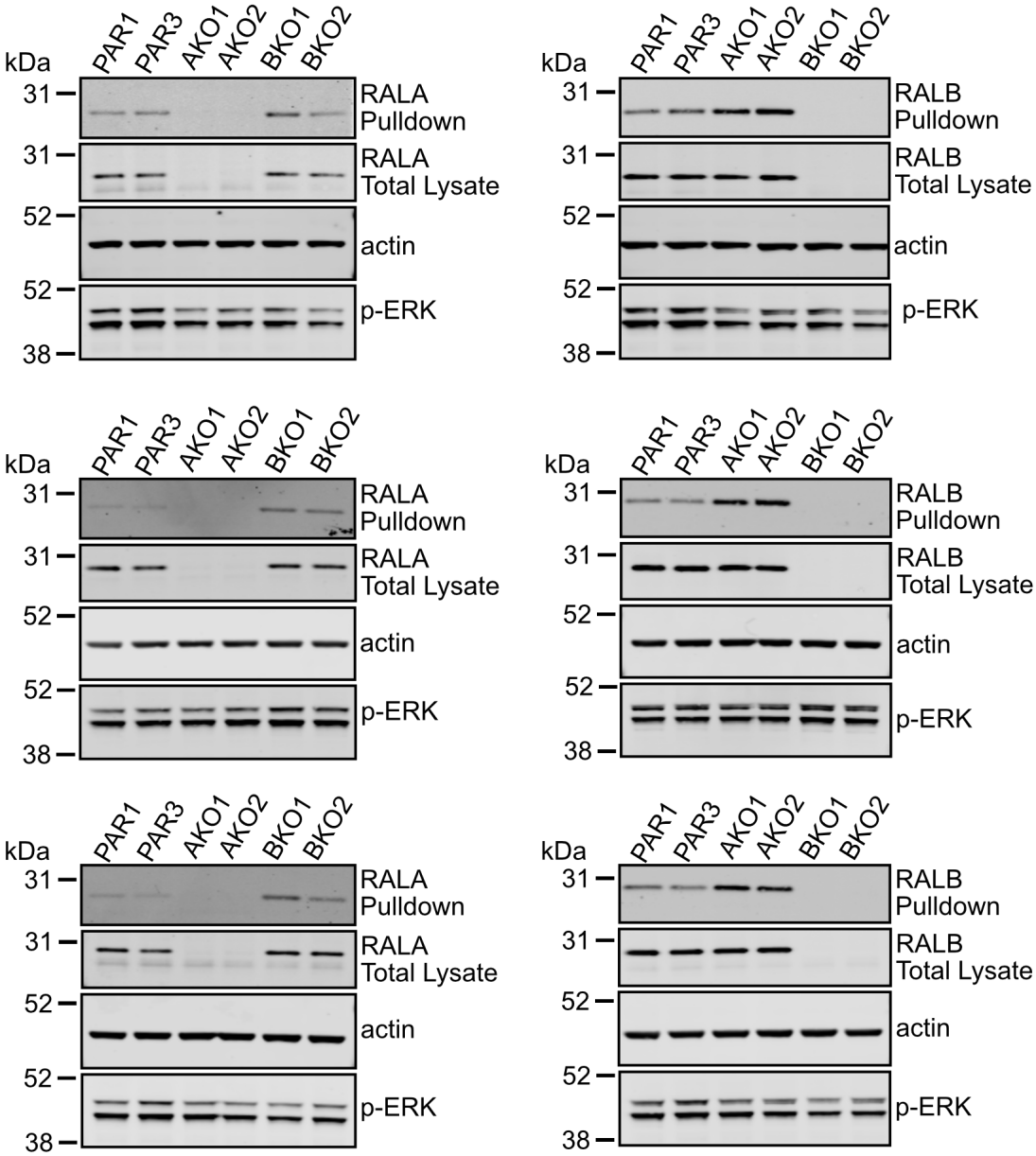
Supplementary Figure 7.8 RAL isoform activity in HCT116 PAR and KO repeats.
Following treatment with 50 ng/mL EGF for 2 minute, HCT116 RALB KO clones, BKO1 and BKO2, displayed a small increase in RALA activity compared to the parental (PAR1 and PAR3). An obvious increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2). Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate.

SW48

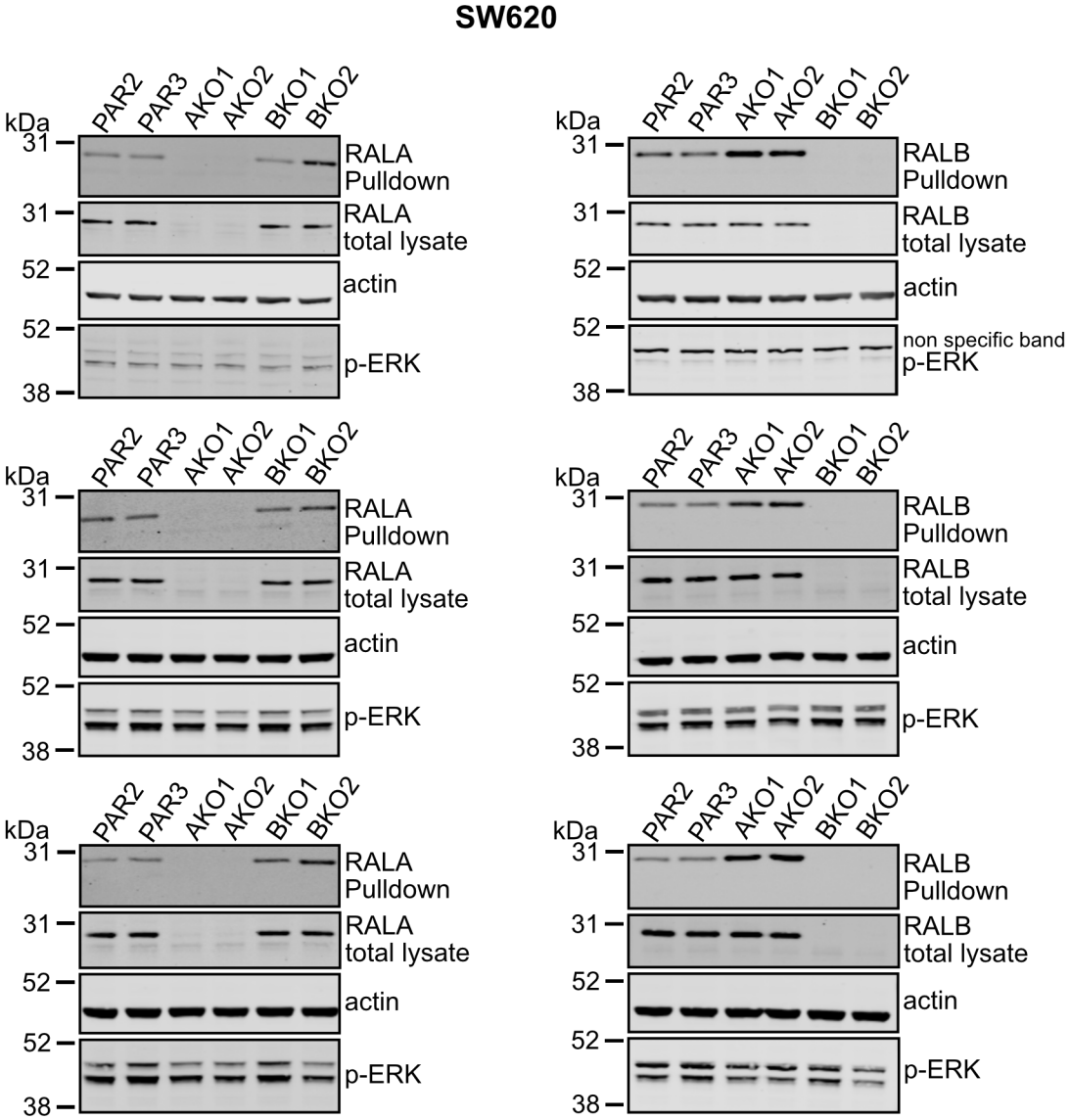


Supplementary Figure 7.9 RAL isoform activity in SW48 PAR and KO repeats.
Following treatment with 50 ng/mL EGF for 2 minutes, a clear increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2) vs PAR2 and PAR3 in SW48. Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate.

SW403



Supplementary Figure 7.10 RAL isoform activity in SW403 PAR and KO repeats.
Following treatment with 50 ng/mL EGF for 2 minute, SW403 RALB KO clones, BKO1 and BKO2, displayed a small increase in RALA activity compared to the parental (PAR1 and PAR3). An obvious increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2). Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate.

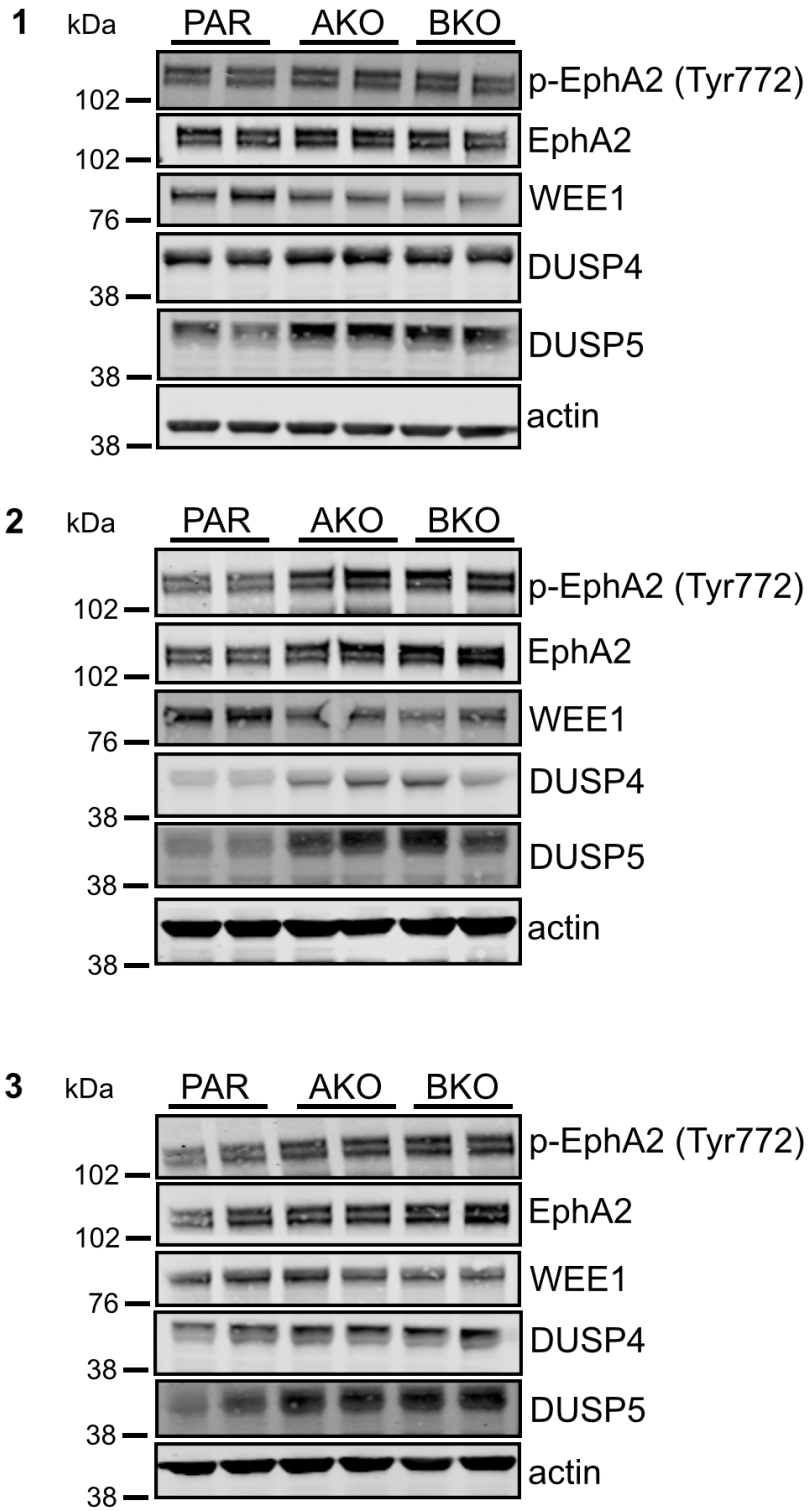


Supplementary Figure 7.11 RAL isoform activity in SW620 PAR and KO repeats. Following treatment with 50 ng/mL EGF for 2 minute, SW620 RALB KO clones, BKO2 in particular, displayed a small increase in RALA activity compared to the parental (PAR1 and PAR3). An obvious increase in RALB activity was observed in both RALA KO clones (AKO1 and AKO2). Whole lysates were prepared, and 400 µg was incubated with GST-RALBD beads which were run on SDS-PAGE along with 10 µg total lysate.

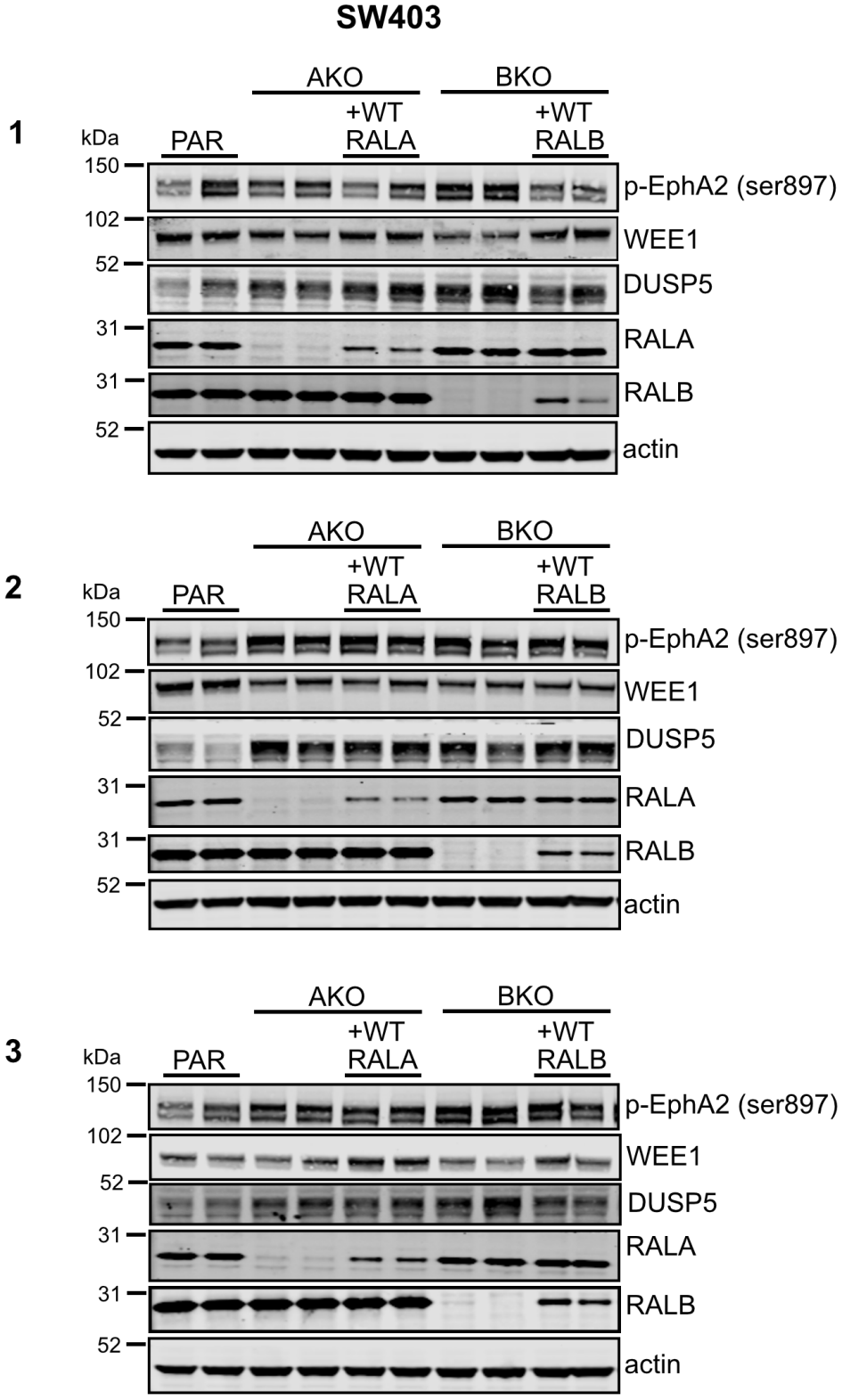
Supplementary Table 7.1. GO analysis for SW403 DEGs collected from RNAseq.

GO code	Description	Gene list
GO:0006260	DNA replication	LIG1, RMI2, MCM2, RNASEH2A, MCM3, MCM4, MCM5, MCM6, RPA1, POLD4, POLE2, CHTF18, RBM14, EGF, ORC1
GO:0001525	angiogenesis	COL18A1, NRP1, TNFRSF12A, NOX1, CSPG4, JAG1, MMRN2, TGFB2, EREG, SRPX2, UNC5B, ID1, PTK2B, DLL4, PLCD3, ADM2, EGF, PLXND1
GO:0000082	G1/S transition of mitotic cell cycle	RPA1, CDKN1A, PLK3, POLE2, PLK2, MCM2, MCM3, MCM4, ORC1, MCM5, MCM6
GO:0000188	inactivation of MAPK activity	DUSP5, DUSP4, DUSP18, DUSP1, RGS3, DUSP8
GO:0008134	transcription factor binding	BCL10, ARHGEF2, CEBPB, FOXA3, FHL2, SMAD3, HES6, IFI16, FOSB, DDIT3, TRIB1, MED19, GATA2, FOS, ID1, ETS1, GPX3, BCL3, PITX2
GO:0016266	O-glycan processing	ST3GAL1, MUC2, ST6GAL1, GALNT6, GALNT5, B3GNT6, GCNT1, MUC4
GO:0031581	hemidesmosome assembly	COL17A1, LAMB3, LAMA3, LAMC2
GO:0006915	apoptotic process	STEAP3, DPF2, KRT20, ZC3H8, UNC5B, PTK2B, PCSK9, INPP5D, PHLDA2, IL1A, PHLDA1, BCL10, BCL2L14, PTPRH, MCM2, DAPK2, DDIT3, WDR92, RNF130, RASSF5, PLK3, RASSF6, BBC3, CSRNP1, GADD45A, DRAM1, PPP1R15A, TRIM39
GO:0007173	EGFR signalling pathway	EREG, PLCG1, PTK2B, GAREM1, HBEGF, AREG, EGF
GO:0034976	response to endoplasmic reticulum stress	CEBPB, WFS1, BBC3, ERN1, TRIB3, FAM129A, PPP1R15A, DDIT3
GO:0008284	positive regulation of cell proliferation	COL18A1, FGF19, NOX1, GAREM1, EFNB2, IL6R, PROX1, PLAC8, TGFB2, PTHLH, AKR1C2, ATF3, EREG, ETS1, PTK2B, CLCF1, ADRA2A, HBEGF, CEACAM6, CHRNA7, LAMC2, AREG, EGF, FOSL1
GO:0009612	response to mechanical stimulus	TXNIP, ETS1, PTK2B, FOSB, IGFBP2, FOSL1, KALRN
GO:0003779	actin binding	CAP2, CNN3, MYO7B, BAIAP2L1, MICAL2, ABLIM3, FSCN1, SPIRE1, FMN1, CORO1A, SYNE1, TNS1, EPB41L1, CGN, GSN, FHOD1
GO:0008285	negative regulation of cell proliferation	COL18A1, KLF10, TIMP2, PROX1, TGFB2, PTHLH, RASSF5, CDKN1A, SERPINE2, EREG, CDKN2B, NME1, ETS1, PTK2B, DLL4, FOSL1, SMARCA2, IL1A, KLF4, NEURL1
GO:0022617	extracellular matrix disassembly	LAMB3, LAMA3, GSN, MMP7, LAMC2, TIMP2, SPP1
GO:0007264	small GTPase mediated signal transduction	RAB3B, RAB7B, RALGDS, RND3, RAB32, ARHGAP32, SH2D3A, RAB30, RAB43, ARL14, RASL10B, RAB14, RAPGEF1, RASD2
GO:0048870	cell motility	SRPX2, ETS1, FSCN1, EPHA2

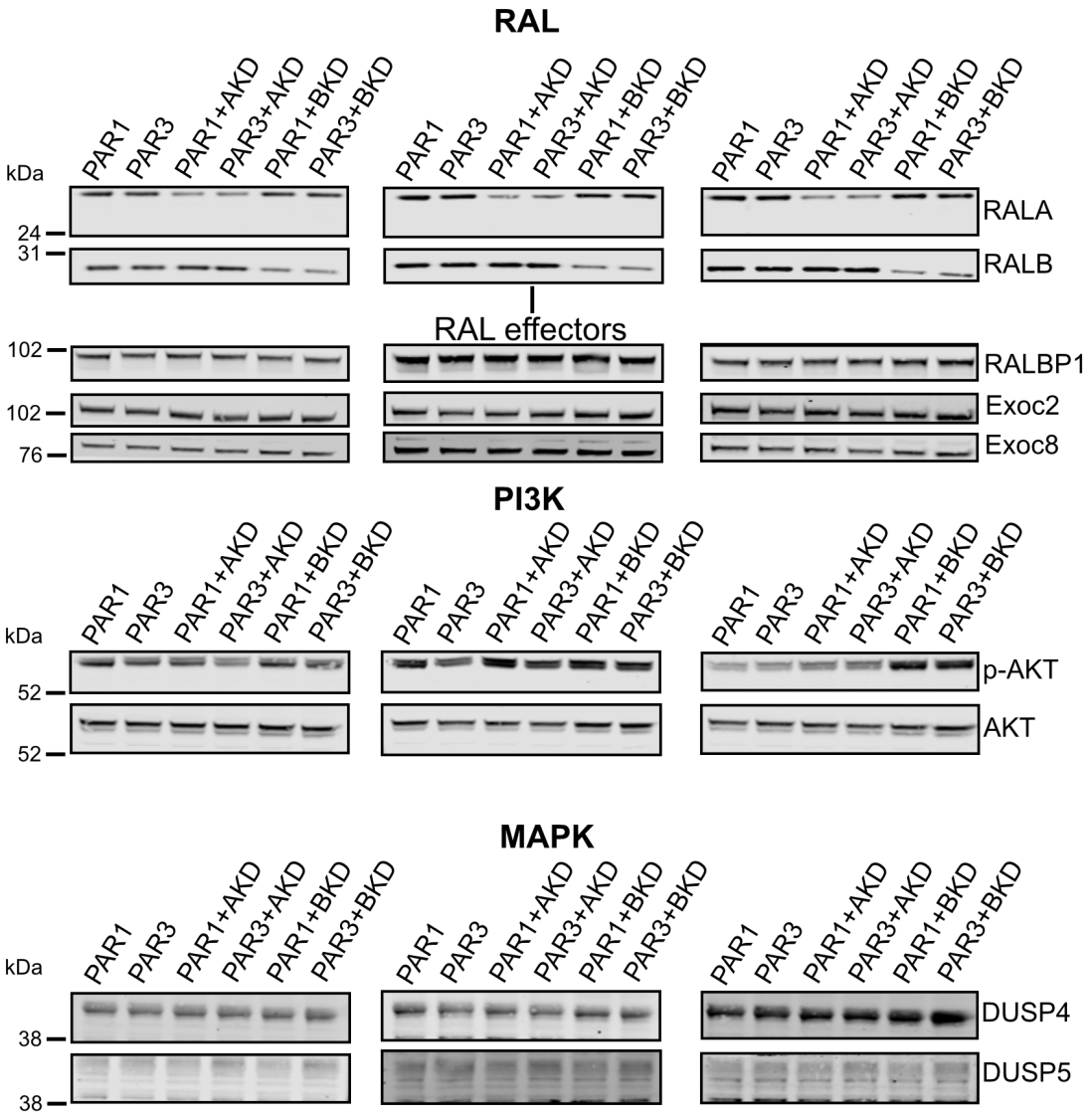
SW403



Supplementary Figure 7.12 Validation of key hits from RNAseq and RPPA data in SW403 KOs repeats.
Immunoblotting of EphA2 and WEE1 confirmed the changes in protein expression detected in the RPPA analysis, with an increase in EphA2 and a decrease in WEE1 levels in the RAL AKO and BKO clones versus parental. An increase in DUSP5 expression was seen in the KO samples, which contrasts with the decrease in transcript expression observed in the RNAseq dataset.

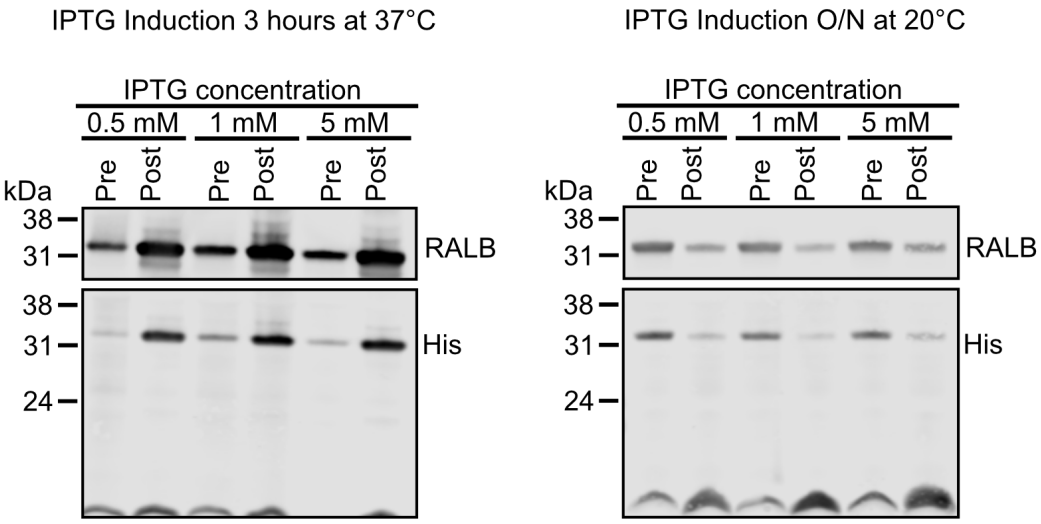


Supplementary Figure 7.13 Validation of key hits from RNAseq and RPPA data in SW403 parental, isoform specific RAL knockout clones and knockout clones stably re-expressing WT RAL repeats.
SW403 RALA knockout clones were transduced with retrovirus to re-express WT RALA and RALB knockout clones were transduced with retrovirus to re-express WT RALB. Expression of WT RAL partially restored WEE1, DUSP5 and p-EphA2 (Ser897) expression to that of the parental clones in some cases. Expression levels of transduced WT RAL do not reach those of the endogenous levels seen in the parental clones.

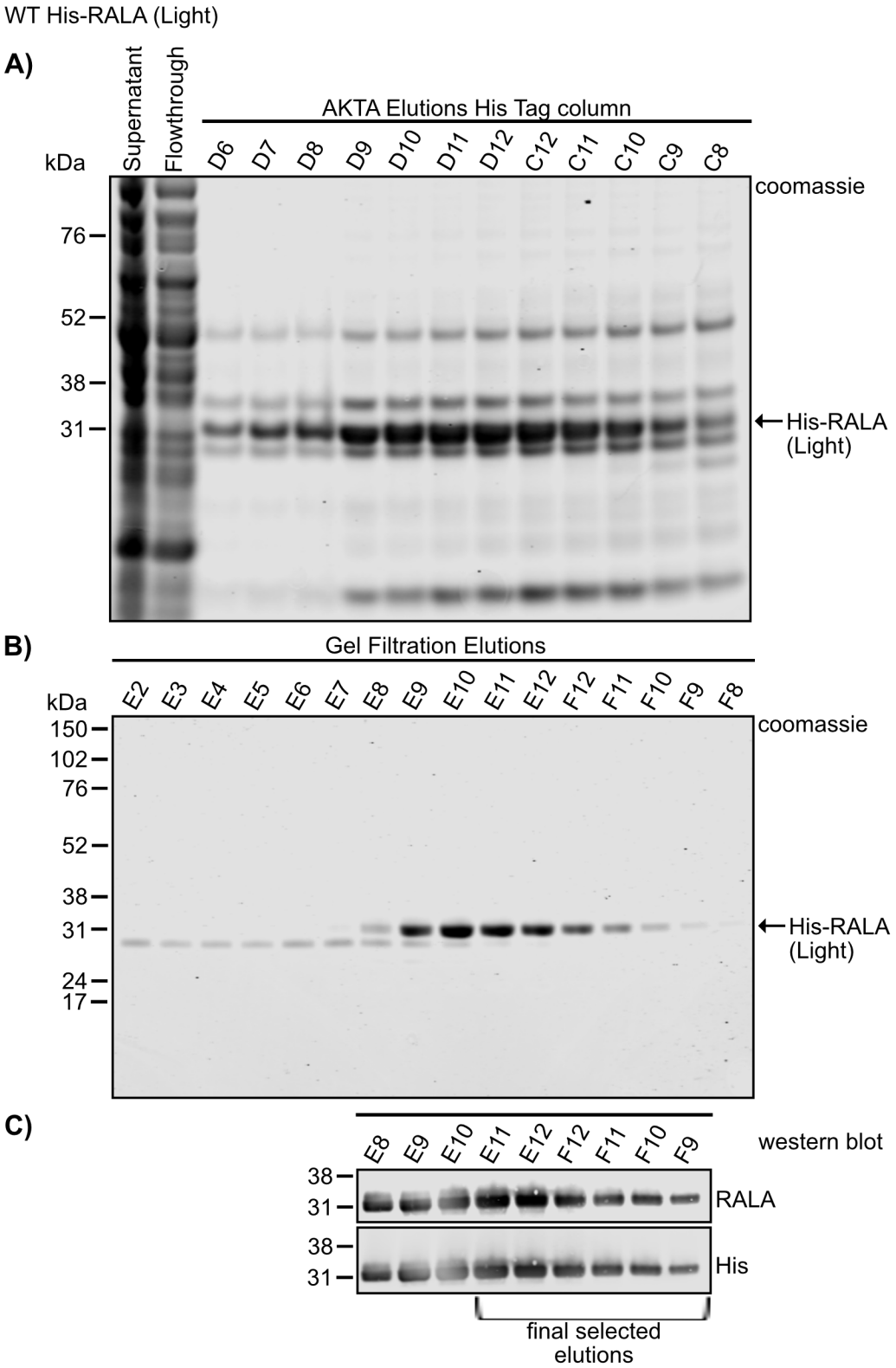


Supplementary Figure 7.14 Expression of known and potential RAL interactors in SW403 PAR clones following isoform specific knockdown.
No expression changes were observed in known RAL effectors following isoform specific knockdown of RAL in SW403 parental clones (PAR1 and PAR3). There were also no clear expression changes in potential hits identified from the RNAseq and RPPA analysis following RAL knockdown in the parental clones.

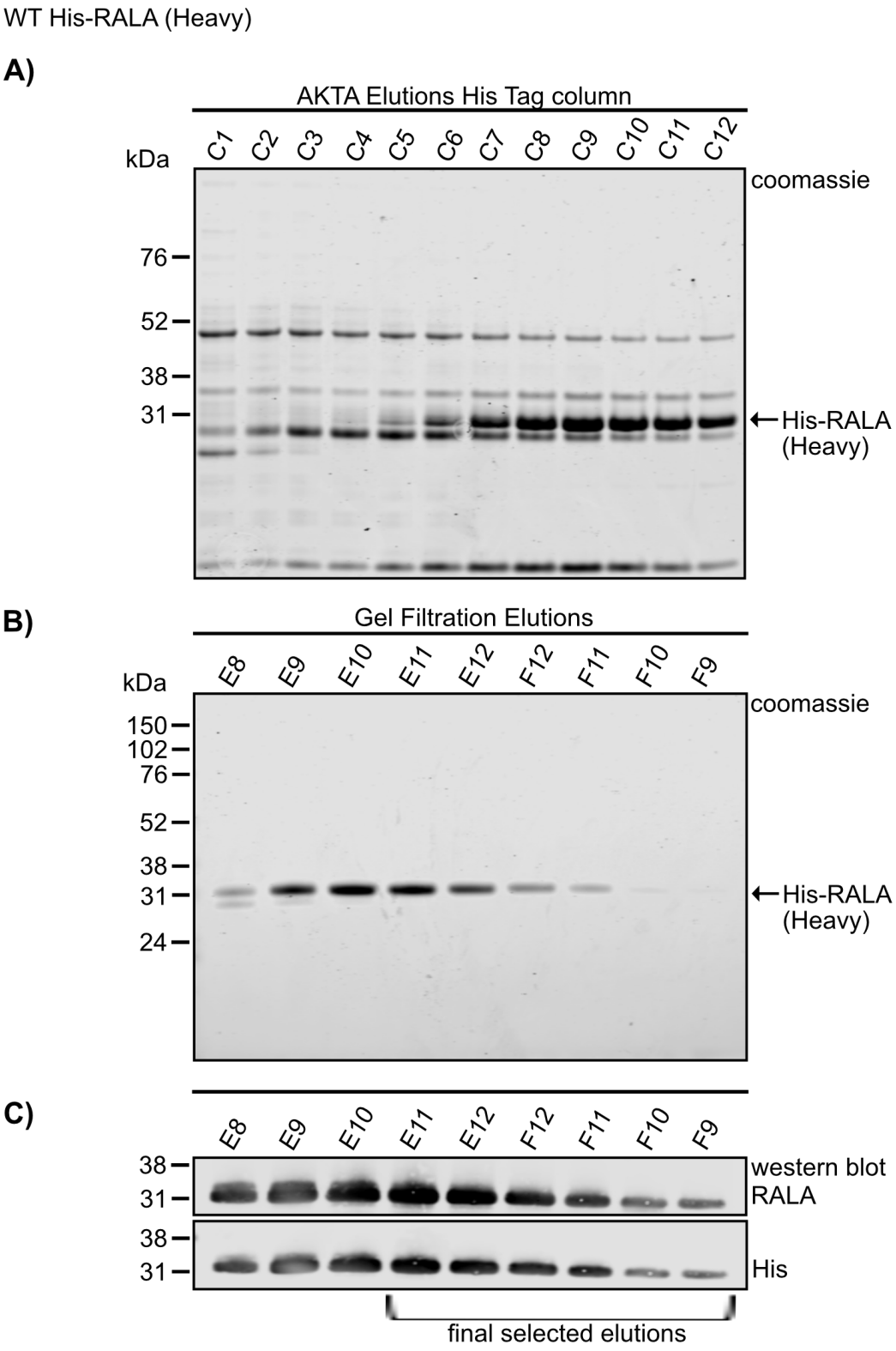
His-RALB



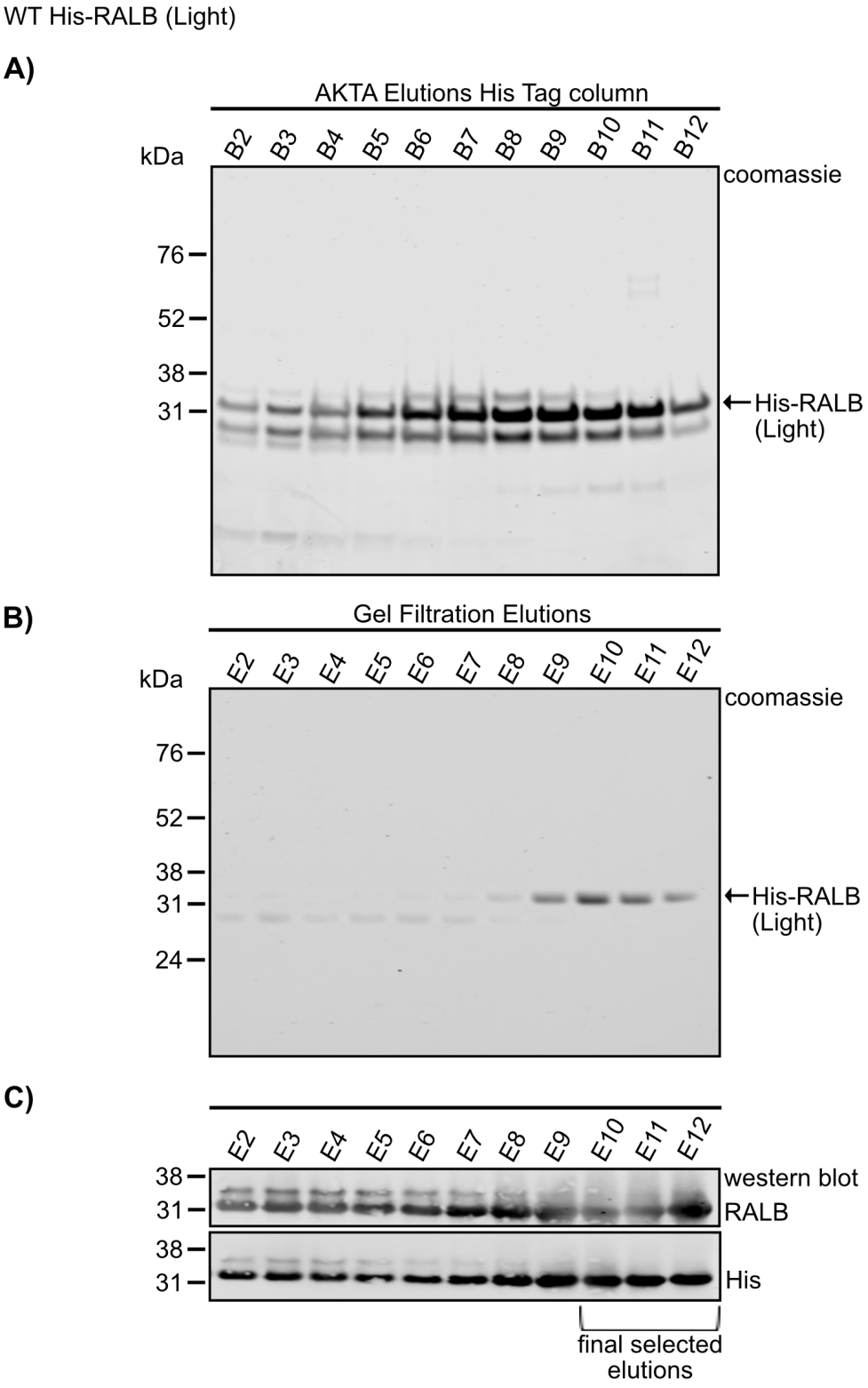
Supplementary Figure 7.15 Test induction of recombinant His-RALB protein.
The optimal induction conditions of His-RalB were determined using increasing IPTG concentrations (0.5-5 mM), combined with either a 3-hour incubation at 37°C or overnight at room temperature. Note 3 hours at the higher temperature is sufficient for His-RalB induction as a longer incubation leads to degradation of the protein.



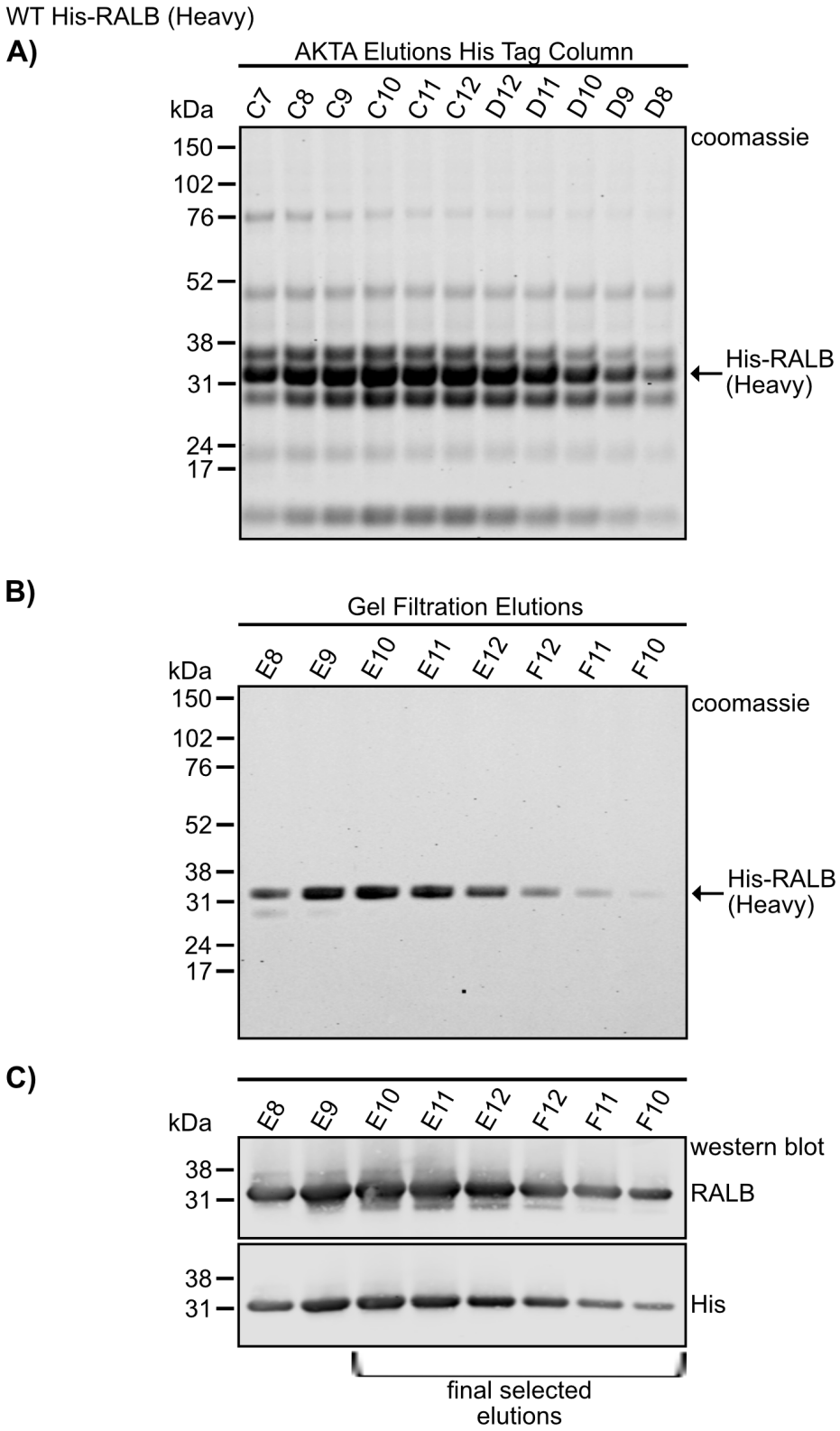
Supplementary Figure 7.16 Protein purification of WT His-RALA (Light) using the AKTA purifier system.
(A) Elutions collected after initial purification on the AKTA purifier system using HisTrap HP following loading of AT713 bacterial supernatant. Induced WT His-RALA (light) protein was present in fractions D6-C8. **(B)** WT His-RALA (light) protein was present in elutions E8-F8 following gel filtration. Fractions E2-E12 show contamination with different protein. **(C)** Western blot of gel filtration elutions E8-F9 confirm the presence of RALA and the co-localisation of the His-tag.



Supplementary Figure 7.17 Protein purification of WT His-RALA (Heavy) using the AKTA purifier system.
(A) Elutions collected after initial purification on the AKTA purifier system using HisTrap HP following loading of AT713 bacterial supernatant. Induced WT His-RALA (heavy) protein was present in fractions C5-C12. **(B)** WT His-RALA (heavy) protein was present in elutions E8-F9 following gel filtration. Fractions E8-E9 show contamination with different protein. **(C)** Western blot of gel filtration elutions E8-F9 confirm the presence of RALA and the co-localisation of the His-tag.

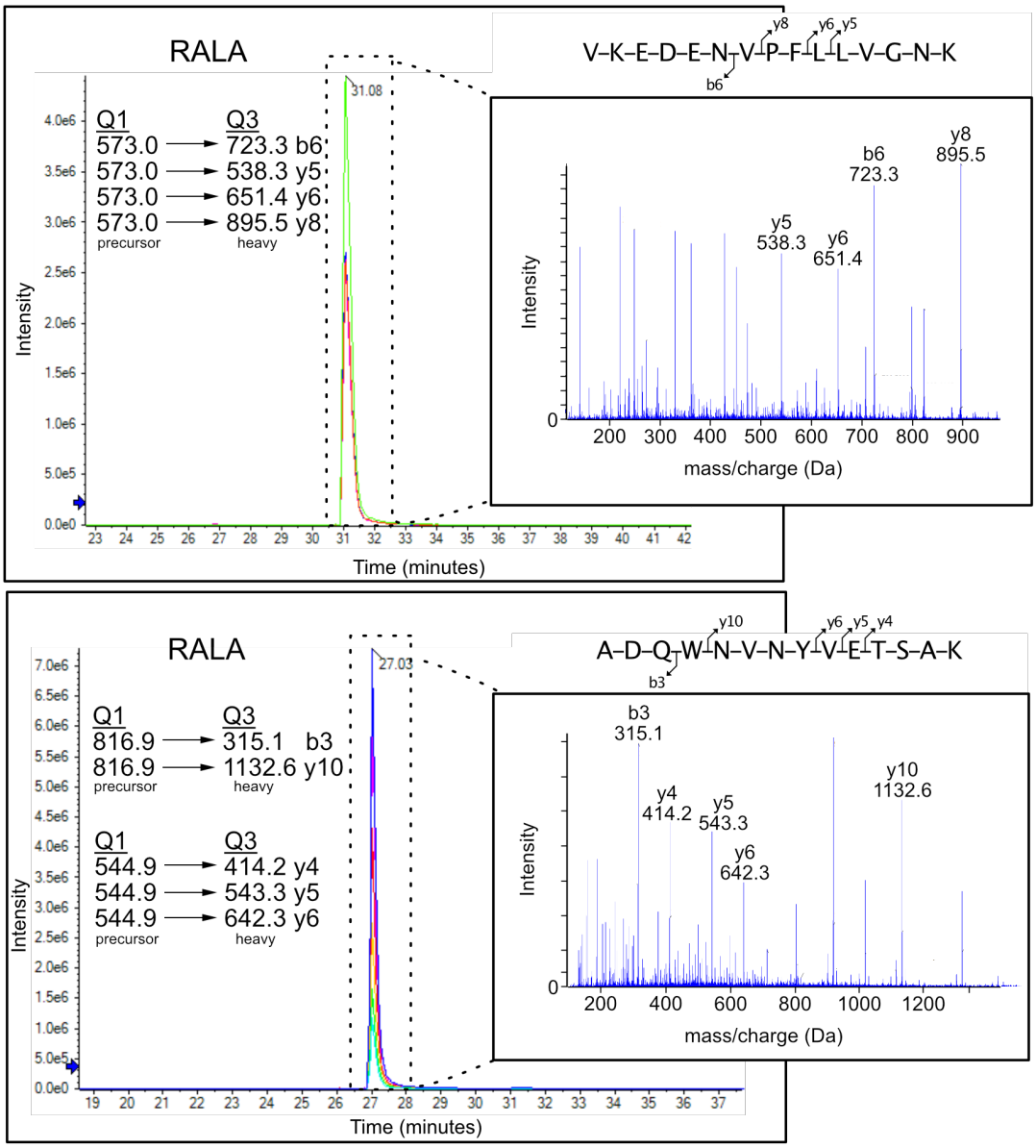


Supplementary Figure 7.18 Protein purification of WT His-RALB (Light) using the AKTA purifier system.
(A) Elutions collected after initial purification on the AKTA purifier system using HisTrap HP following loading of AT713 bacterial supernatant. Induced WT His-RALB (light) protein was present in fractions B2-B12. **(B)** WT His-RALB (light) protein was present in elutions E2-E12 following gel filtration. Fractions E2-E9 show contamination with different protein. **(C)** Western blot of gel filtration elutions E2-F9 confirm the presence of RALB and the co-localisation of the His-tag.



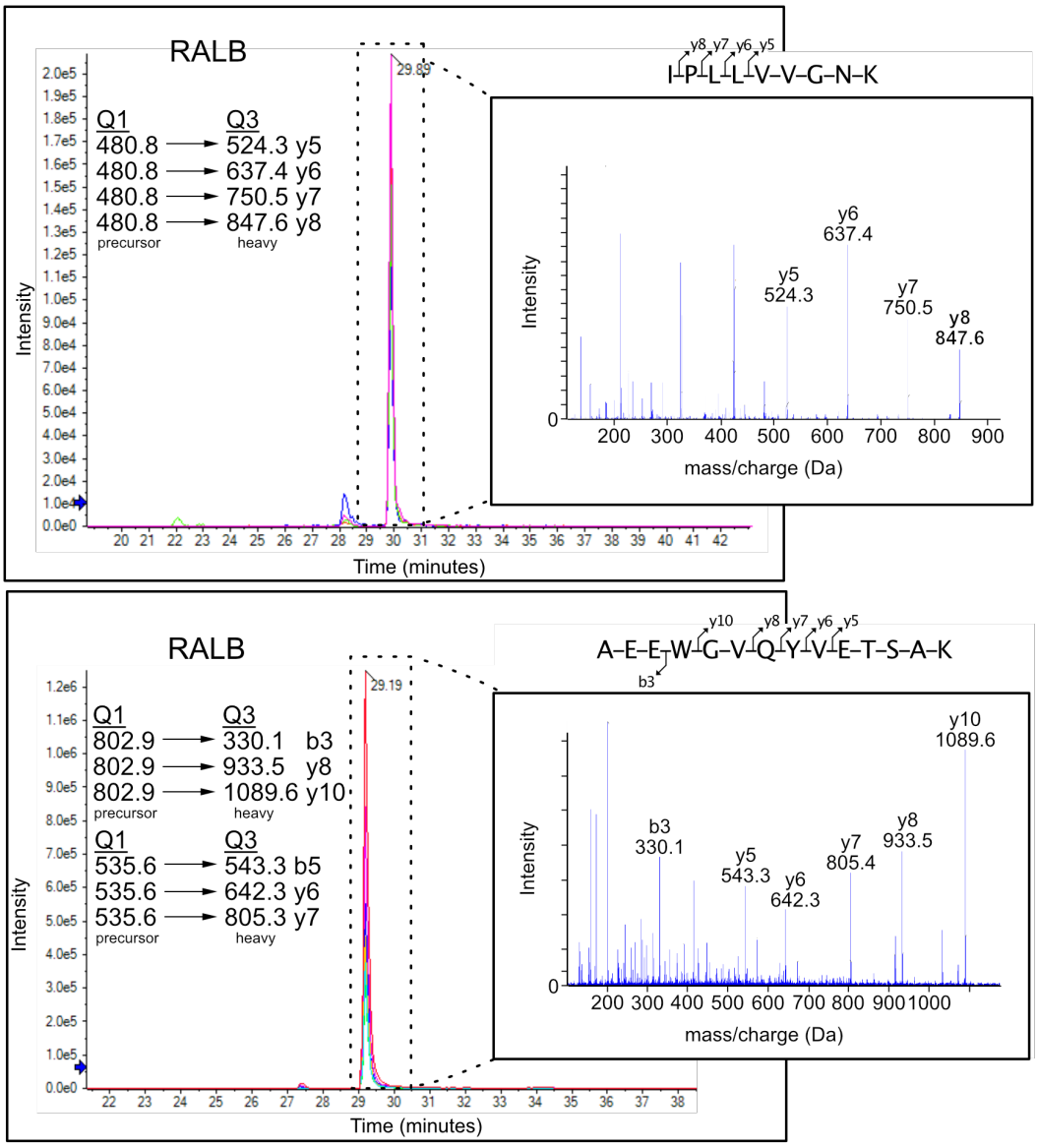
Supplementary Figure 7.19 Protein purification of WT His-RALB (Heavy) using the AKTA purifier system.

(A) Elutions collected after initial purification on the AKTA purifier system using HisTrap HP following loading of AT713 bacterial supernatant. Induced His-RALB (heavy) protein was present in fractions C7-D8. **(B)** His-RALB (heavy) protein was present in elutions E8-F10 following gel filtration. Fractions E8-E9 show contamination with different protein. **(C)** Western blot of gel filtration elutions E8-F10 confirm the presence of RALB and the co-localisation of the His-tag.



		PRECURSOR ION (m/z)		PRODUCT ION (m/z)		
PEPTIDE	CHARGE	K0 R0	K8 R10	K0 R0	K8 R10	ION
VKEDENVPFLLVGNK	+3	567.6	573.0	715.3	723.3	b6
				530.3	538.3	y5
				643.4	651.4	y6
				887.5	895.5	y8
ADQWNVNYVETSAK	+2	812.9	816.9	315.1	315.1	b3
				1124.6	1132.6	y10
	+3	542.3	544.9	406.2	414.2	y4
				535.3	543.3	y5
				634.3	642.3	y6

Supplementary Figure 7.20 MS/MS transition spectra for RALA peptides.
100 ng of His-tagged RAL isoforms were subjected to LC-MS/MS analysis for detection of candidate peptides. Both RALA specific peptides were detected. Despite VKEDENVPFLLVGNK including a miscleavage, the more consistent readings meant it was selected as the final peptide for future calculations.



		PRECURSOR ION (m/z)		PRODUCT ION (m/z)		
PEPTIDE	CHARGE	K0 R0	K8 R10	K0 R0	K8 R10	ION
IPLLVVGNK	+2	476.8	480.8	516.3	524.3	y5
				629.4	637.4	y6
				742.5	750.5	y7
				839.5	847.6	y8
AEEWGVQYVETSAK	+2	798.8	802.9	330.1	330.1	b3
				925.5	933.5	y8
				1081.6	1089.6	y10
	+3	532.9	535.6	535.3	543.3	y5
				634.3	642.3	y6
				797.4	805.3	y7

Supplementary Figure 7.21 MS/MS transition spectra for RALB peptides.
100 ng of His-tagged RAL isoforms were subjected to LC-MS/MS analysis for detection of candidate peptides. Both RALB specific peptides were detected at this concentration. However, when repeated at lower concentrations required for sample processing, peptide IPLLVVGNK was absent.

References

Adhikari, H. and Counter, C. M. (2018) 'Interrogating the protein interactomes of RAS isoforms identifies PIP5K1A as a KRAS-specific vulnerability', *Nat Commun,* 9(1), pp. 3646.
Aebersold, R. and Mann, M. (2003) 'Mass spectrometry-based proteomics', *Nature,* 422(6928), pp. 198-207.
Agapova, L. S., Volodina, J. L., Chumakov, P. M. and Kopnin, B. P. (2004) 'Activation of Ras-Ral pathway attenuates p53-independent DNA damage G2 checkpoint', *J Biol Chem,* 279(35), pp. 36382-9.
Aguirre, A. J., Meyers, R. M., Weir, B. A., Vazquez, F., Zhang, C. Z., Ben-David, U., Cook, A., Ha, G., Harrington, W. F., Doshi, M. B., Kost-Alimova, M., Gill, S., Xu, H., Ali, L. D., Jiang, G., Pantel, S., Lee, Y., Goodale, A., Cherniack, A. D., Oh, C., Kryukov, G., Cowley, G. S., Garraway, L. A., Stegmaier, K., Roberts, C. W., Golub, T. R., Meyerson, M., Root, D. E., Tsherniak, A. and Hahn, W. C. (2016) 'Genomic Copy Number Dictates a Gene-Independent Cell Response to CRISPR/Cas9 Targeting', *Cancer Discov,* 6(8), pp. 914-29.
Aguirre-Ghiso, J. A., Frankel, P., Farias, E. F., Lu, Z., Jiang, H., Olsen, A., Feig, L. A., de Kier Joffe, E. B. and Foster, D. A. (1999) 'RalA requirement for v-Src- and v-Ras-induced tumorigenicity and overproduction of urokinase-type plasminogen activator: involvement of metalloproteases', *Oncogene,* 18(33), pp. 4718-25.
Albright, C. F., Giddings, B. W., Liu, J., Vito, M. and Weinberg, R. A. (1993) 'Characterization of a guanine nucleotide dissociation stimulator for a ras-related GTPase', *EMBO J,* 12(1), pp. 339-47.
Aldridge, B. B., Burke, J. M., Lauffenburger, D. A. and Sorger, P. K. (2006) 'Physicochemical modelling of cell signalling pathways', *Nat Cell Biol,* 8(11), pp. 1195-203.
Anderson, L. and Hunter, C. L. (2006) 'Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins', *Mol Cell Proteomics,* 5(4), pp. 573-88.
Arkin, M. R., Tang, Y. and Wells, J. A. (2014) 'Small-molecule inhibitors of protein-protein interactions: progressing toward the reality', *Chem Biol,* 21(9), pp. 1102-14.
Audic, S. and Claverie, J. M. (1997) 'The significance of digital gene expression profiles', *Genome Res,* 7(10), pp. 986-95.
Baldelli, E., Calvert, V., Hodge, A., VanMeter, A., Petricoin, E. F. and Pierobon, M. (2017) 'Reverse Phase Protein Microarrays', *Methods Mol Biol,* 1606, pp. 149-169.
Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. and Horvath, P. (2007) 'CRISPR provides acquired resistance against viruses in prokaryotes', *Science,* 315(5819), pp. 1709-12.
Beel, S., Kolloch, L., Apken, L. H., Jürgens, L., Bolle, A., Sudhof, N., Ghosh, S., Wardelmann, E., Meisterernst, M., Steinestel, K. and Oeckinghaus, A. (2020) 'κB-Ras and Ral GTPases regulate acinar to ductal metaplasia during pancreatic adenocarcinoma development and pancreatitis', *Nat Commun,* 11(1), pp. 3409.
Behan, F. M., Iorio, F., Picco, G., Gonçalves, E., Beaver, C. M., Migliardi, G., Santos, R., Rao, Y., Sassi, F., Pinnelli, M., Ansari, R., Harper, S., Jackson, D. A., McRae, R., Pooley, R., Wilkinson, P., van der Meer, D., Dow, D., Buser-

Doepner, C., Bertotti, A., Trusolino, L., Stronach, E. A., Saez-Rodriguez, J., Yusa, K. and Garnett, M. J. (2019) 'Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens', *Nature,* 568(7753), pp. 511-516.
Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A. and Bonas, U. (2009) 'Breaking the code of DNA binding specificity of TAL-type III effectors', *Science,* 326(5959), pp. 1509-12.
Bodemann, B. O., Orvedahl, A., Cheng, T., Ram, R. R., Ou, Y. H., Formstecher, E., Maiti, M., Hazelett, C. C., Wauson, E. M., Balakireva, M., Camonis, J. H., Yeaman, C., Levine, B. and White, M. A. (2011) 'RalB and the exocyst mediate the cellular starvation response by direct activation of autophagosome assembly', *Cell,* 144(2), pp. 253-67.
Bodempudi, V., Yamoutpoor, F., Pan, W., Dudek, A. Z., Esfandyari, T., Piedra, M., Babovick-Vuksanovic, D., Woo, R. A., Mautner, V. F., Kluwe, L., Clapp, D. W., De Vries, G. H., Thomas, S. L., Kurtz, A., Parada, L. F. and Farassati, F. (2009) 'Ral overactivation in malignant peripheral nerve sheath tumors', *Mol Cell Biol,* 29(14), pp. 3964-74.
Boellner, S. and Becker, K. F. (2015) 'Reverse Phase Protein Arrays-Quantitative Assessment of Multiple Biomarkers in Biopsies for Clinical Use', *Microarrays (Basel),* 4(2), pp. 98-114.
Boettcher, M. and McManus, M. T. (2015) 'Choosing the Right Tool for the Job: RNAi, TALEN, or CRISPR', *Mol Cell,* 58(4), pp. 575-85.
Bogdanove, A. J. and Voytas, D. F. (2011) 'TAL effectors: customizable proteins for DNA targeting', *Science,* 333(6051), pp. 1843-6.
Bordbar, A., Monk, J. M., King, Z. A. and Palsson, B. O. (2014) 'Constraint-based models predict metabolic and associated cellular functions', *Nat Rev Genet,* 15(2), pp. 107-20.
Brun, V., Masselon, C., Garin, J. and Dupuis, A. (2009) 'Isotope dilution strategies for absolute quantitative proteomics', *J Proteomics,* 72(5), pp. 740-9.
Buffet, C., Catelli, M. G., Hecale-Perlemoine, K., Bricaire, L., Garcia, C., Gallet-Dierick, A., Rodriguez, S., Cormier, F. and Groussin, L. (2015) 'Dual Specificity Phosphatase 5, a Specific Negative Regulator of ERK Signaling, Is Induced by Serum Response Factor and Elk-1 Transcription Factor', *PLoS One,* 10(12), pp. e0145484.
Bum-Erdene, K., Liu, D., Gonzalez-Gutierrez, G., Ghozayel, M. K., Xu, D. and Meroueh, S. O. (2020) 'Small-molecule covalent bond formation at tyrosine creates a binding site and inhibits activation of Ral GTPases', *Proc Natl Acad Sci U S A,* 117(13), pp. 7131-7139.
Camonis, J. H. and White, M. A. (2005) 'Ral GTPases: corrupting the exocyst in cancer cells', *Trends Cell Biol,* 15(6), pp. 327-32.
Cantor, S. B., Urano, T. and Feig, L. A. (1995) 'Identification and characterization of Ral-binding protein 1, a potential downstream target of Ral GTPases', *Mol Cell Biol,* 15(8), pp. 4578-84.
Capecchi, M. R. (1989) 'Altering the genome by homologous recombination', *Science,* 244(4910), pp. 1288-92.
Capecchi, M. R. (2005) 'Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century', *Nat Rev Genet,* 6(6), pp. 507-12.
Carroll, D. (2011) 'Genome engineering with zinc-finger nucleases', *Genetics,* 188(4), pp. 773-82.

Carthew, R. W. and Sontheimer, E. J. (2009) 'Origins and Mechanisms of miRNAs and siRNAs', *Cell,* 136(4), pp. 642-55.
Caunt, C. J. and Keyse, S. M. (2013) 'Dual-specificity MAP kinase phosphatases (MKPs): shaping the outcome of MAP kinase signalling', *FEBS J,* 280(2), pp. 489-504.
Chapman, J. R., Taylor, M. R. and Boulton, S. J. (2012) 'Playing the end game: DNA double-strand break repair pathway choice', *Mol Cell,* 47(4), pp. 497-510.
Chapman, K. and Hall, D. J. (2005) 'Quantifying the activation of the small molecular weight G-protein ras', *Biochem Mol Biol Educ,* 33(1), pp. 22-7.
Chardin, P. and Tavitian, A. (1986) 'The ral gene: a new ras related gene isolated by the use of a synthetic probe', *EMBO J,* 5(9), pp. 2203-8.
Chardin, P. and Tavitian, A. (1989) 'Coding sequences of human ralA and ralB cDNAs', *Nucleic Acids Res,* 17(11), pp. 4380.
Cheadle, J. P., Krawczak, M., Thomas, M. W., Hodges, A. K., Al-Tassan, N., Fleming, N. and Sampson, J. R. (2002) 'Different combinations of biallelic APC mutation confer different growth advantages in colorectal tumours', *Cancer Res,* 62(2), pp. 363-6.
Chen, F., Pruett-Miller, S. M., Huang, Y., Gjoka, M., Duda, K., Taunton, J., Collingwood, T. N., Frodin, M. and Davis, G. D. (2011a) 'High-frequency genome editing using ssDNA oligonucleotides with zinc-finger nucleases', *Nat Methods,* 8(9), pp. 753-5.
Chen, L., Morrow, J. K., Tran, H. T., Phatak, S. S., Du-Cuny, L. and Zhang, S. (2012) 'From laptop to benchtop to bedside: structure-based drug design on protein targets', *Curr Pharm Des,* 18(9), pp. 1217-39.
Chen, S., Sanjana, N. E., Zheng, K., Shalem, O., Lee, K., Shi, X., Scott, D. A., Song, J., Pan, J. Q., Weissleder, R., Lee, H., Zhang, F. and Sharp, P. A. (2015) 'Genome-wide CRISPR screen in a mouse model of tumor growth and metastasis', *Cell,* 160(6), pp. 1246-60.
Chen, W., Smeekens, J. M. and Wu, R. (2016) 'Systematic study of the dynamics and half-lives of newly synthesized proteins in human cells', *Chem Sci,* 7(2), pp. 1393-1400.
Chen, X. W., Leto, D., Xiong, T., Yu, G., Cheng, A., Decker, S. and Saltiel, A. R. (2011b) 'A Ral GAP complex links PI 3-kinase/Akt signaling to RalA activation in insulin action', *Mol Biol Cell,* 22(1), pp. 141-52.
Chesik, D., De Keyser, J., Bron, R. and Fuhler, G. M. (2010) 'Insulin-like growth factor binding protein-1 activates integrin-mediated intracellular signaling and migration in oligodendrocytes', *J Neurochem,* 113(5), pp. 1319-30.
Chew, W. L., Tabebordbar, M., Cheng, J. K., Mali, P., Wu, E. Y., Ng, A. H., Zhu, K., Wagers, A. J. and Church, G. M. (2016) 'A multifunctional AAV-CRISPR-Cas9 and its host response', *Nat Methods,* 13(10), pp. 868-74.
Chien, Y., Kim, S., Bumeister, R., Loo, Y. M., Kwon, S. W., Johnson, C. L., Balakireva, M. G., Romeo, Y., Kopelovich, L., Gale, M., Yeaman, C., Camonis, J. H., Zhao, Y. and White, M. A. (2006) 'RalB GTPase-mediated activation of the IkappaB family kinase TBK1 couples innate immune signaling to tumor cell survival', *Cell,* 127(1), pp. 157-70.
Chien, Y. and White, M. A. (2003) 'RAL GTPases are linchpin modulators of human tumour-cell proliferation and survival', *EMBO Rep,* 4(8), pp. 800-6.

Cho, S. W., Kim, S., Kim, J. M. and Kim, J. S. (2013) 'Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease', *Nat Biotechnol,* 31(3), pp. 230-2.
Chu, V. T., Weber, T., Wefers, B., Wurst, W., Sander, S., Rajewsky, K. and Kühn, R. (2015) 'Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells', *Nat Biotechnol,* 33(5), pp. 543-8.
Cong, L., Ran, F. A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P. D., Wu, X., Jiang, W., Marraffini, L. A. and Zhang, F. (2013) 'Multiplex genome engineering using CRISPR/Cas systems', *Science,* 339(6121), pp. 819-23.
Cowley, S., Paterson, H., Kemp, P. and Marshall, C. J. (1994) 'Activation of MAP kinase kinase is necessary and sufficient for PC12 differentiation and for transformation of NIH 3T3 cells', *Cell,* 77(6), pp. 841-52.
Cox, A. D. and Der, C. J. (2010) 'Ras history: The saga continues', *Small GTPases,* 1(1), pp. 2-27.
Cox, A. D., Fesik, S. W., Kimmelman, A. C., Luo, J. and Der, C. J. (2014) 'Drugging the undruggable RAS: Mission possible?', *Nat Rev Drug Discov,* 13(11), pp. 828-51.
Cullis, D. N., Philip, B., Baleja, J. D. and Feig, L. A. (2002) 'Rab11-FIP2, an adaptor protein connecting cellular components involved in internalization and recycling of epidermal growth factor receptors', *J Biol Chem,* 277(51), pp. 49158-66.
D'Adamo, D. R., Novick, S., Kahn, J. M., Leonardi, P. and Pellicer, A. (1997) 'rsc: a novel oncogene with structural and functional homology with the gene family of exchange factors for Ral', *Oncogene,* 14(11), pp. 1295-305.
de Bruyn, K. M., de Rooij, J., Wolthuis, R. M., Rehmann, H., Wesenbeek, J., Cool, R. H., Wittinghofer, A. H. and Bos, J. L. (2000) 'RalGEF2, a pleckstrin homology domain containing guanine nucleotide exchange factor for Ral', *J Biol Chem,* 275(38), pp. 29761-6.
de Gorter, D. J., Vos, J. C., Pals, S. T. and Spaargaren, M. (2007) 'The B cell antigen receptor controls AP-1 and NFAT activity through Ras-mediated activation of Ral', *J Immunol,* 178(3), pp. 1405-14.
de Ruiter, N. D., Wolthuis, R. M., van Dam, H., Burgering, B. M. and Bos, J. L. (2000) 'Ras-dependent regulation of c-Jun phosphorylation is mediated by the Ral guanine nucleotide exchange factor-Ral pathway', *Mol Cell Biol,* 20(22), pp. 8480-8.
Der, C. J., Finkel, T. and Cooper, G. M. (1986) 'Biological and biochemical properties of human rasH genes mutated at codon 61', *Cell,* 44(1), pp. 167-76.
Deutsch, E. W., Lam, H. and Aebersold, R. (2008) 'PeptideAtlas: a resource for target selection for emerging targeted proteomics workflows', *EMBO Rep,* 9(5), pp. 429-34.
Ding, Y., Li, H., Chen, L. L. and Xie, K. (2016) 'Recent Advances in Genome Editing Using CRISPR/Cas9', *Front Plant Sci,* 7, pp. 703.
Domon, B. and Aebersold, R. (2006) 'Mass spectrometry and protein analysis', *Science,* 312(5771), pp. 212-7.
Draper, B. K., Komurasaki, T., Davidson, M. K. and Nanney, L. B. (2003) 'Epiregulin is more potent than EGF or TGFalpha in promoting in vitro wound closure due to enhanced ERK/MAPK activation', *J Cell Biochem,* 89(6), pp. 1126-37.

Drosten, M., Dhawahir, A., Sum, E. Y., Urosevic, J., Lechuga, C. G., Esteban, L. M., Castellano, E., Guerra, C., Santos, E. and Barbacid, M. (2010) 'Genetic analysis of Ras signalling pathways in cell proliferation, migration and survival', *EMBO J,* 29(6), pp. 1091-104.
Dunne, P. D., Dasgupta, S., Blayney, J. K., McArt, D. G., Redmond, K. L., Weir, J. A., Bradley, C. A., Sasazuki, T., Shirasawa, S., Wang, T., Srivastava, S., Ong, C. W., Arthur, K., Salto-Tellez, M., Wilson, R. H., Johnston, P. G. and Van Schaeybroeck, S. (2016) 'EphA2 Expression Is a Key Driver of Migration and Invasion and a Poor Prognostic Marker in Colorectal Cancer', *Clin Cancer Res,* 22(1), pp. 230-242.
Dupuis, A., Hennekinne, J. A., Garin, J. and Brun, V. (2008) 'Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks', *Proteomics,* 8(22), pp. 4633-6.
Eden, E., Geva-Zatorsky, N., Issaeva, I., Cohen, A., Dekel, E., Danon, T., Cohen, L., Mayo, A. and Alon, U. (2011) 'Proteome half-life dynamics in living human cells', *Science,* 331(6018), pp. 764-8.
Emkey, R., Freedman, S. and Feig, L. A. (1991) 'Characterization of a GTPase-activating protein for the Ras-related Ral protein', *J Biol Chem,* 266(15), pp. 9703-6.
Ezzeldin, M., Borrego-Diaz, E., Taha, M., Esfandyari, T., Wise, A. L., Peng, W., Rouyanian, A., Asvadi Kermani, A., Soleimani, M., Patrad, E., Lialyte, K., Wang, K., Williamson, S., Abdulkarim, B., Olyaee, M. and Farassati, F. (2014) 'RalA signaling pathway as a therapeutic target in hepatocellular carcinoma (HCC)', *Mol Oncol,* 8(5), pp. 1043-53.
Falsetti, S. C., Wang, D. A., Peng, H., Carrico, D., Cox, A. D., Der, C. J., Hamilton, A. D. and Sebti, S. M. (2007) 'Geranylgeranyltransferase I inhibitors target RalB to inhibit anchorage-dependent growth and induce apoptosis and RalA to inhibit anchorage-independent growth', *Mol Cell Biol,* 27(22), pp. 8003-14.
Feig, L. A. (2003) 'Ral-GTPases: approaching their 15 minutes of fame', *Trends Cell Biol,* 13(8), pp. 419-25.
Fenwick, R. B., Campbell, L. J., Rajasekar, K., Prasannan, S., Nietlispach, D., Camonis, J., Owen, D. and Mott, H. R. (2010) 'The RalB-RLIP76 complex reveals a novel mode of ral-effector interaction', *Structure,* 18(8), pp. 985-95.
Fernández, R. M., Ruiz-Miró, M., Dolcet, X., Aldea, M. and Garí, E. (2011) 'Cyclin D1 interacts and collaborates with Ral GTPases enhancing cell detachment and motility', *Oncogene,* 30(16), pp. 1936-46.
Ferro, E. and Trabalzini, L. (2010) 'RalGDS family members couple Ras to Ral signalling and that's not all', *Cell Signal,* 22(12), pp. 1804-10.
Fineran, P. C. and Charpentier, E. (2012) 'Memory of viral infections by CRISPR-Cas adaptive immune systems: acquisition of new information', *Virology,* 434(2), pp. 202-9.
Fodde, R. (2002) 'The APC gene in colorectal cancer', *Eur J Cancer,* 38(7), pp. 867-71.
Franceschini, A., Meier, R., Casanova, A., Kreibich, S., Daga, N., Andritschke, D., Dilling, S., Rämö, P., Emmenlauer, M., Kaufmann, A., Conde-Álvarez, R., Low, S. H., Pelkmans, L., Helenius, A., Hardt, W. D., Dehio, C. and von Mering, C. (2014) 'Specific inhibition of diverse pathogens in human cells by synthetic microRNA-like oligonucleotides inferred from RNAi screens', *Proc Natl Acad Sci U S A,* 111(12), pp. 4548-53.

Frankel, P., Aronheim, A., Kavanagh, E., Balda, M. S., Matter, K., Bunney, T. D. and Marshall, C. J. (2005) 'RalA interacts with ZONAB in a cell density-dependent manner and regulates its transcriptional activity', *EMBO J,* 24(1), pp. 54-62.
Frische, E. W., Pellis-van Berkel, W., van Haaften, G., Cuppen, E., Plasterk, R. H., Tijsterman, M., Bos, J. L. and Zwartkruis, F. J. (2007) 'RAP-1 and the RAL-1/exocyst pathway coordinate hypodermal cell organization in Caenorhabditis elegans', *EMBO J,* 26(24), pp. 5083-92.
Fu, Y., Foden, J. A., Khayter, C., Maeder, M. L., Reyon, D., Joung, J. K. and Sander, J. D. (2013) 'High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells', *Nat Biotechnol,* 31(9), pp. 822-6.
Fukai, S., Matern, H. T., Jagath, J. R., Scheller, R. H. and Brunger, A. T. (2003) 'Structural basis of the interaction between RalA and Sec5, a subunit of the sec6/8 complex', *EMBO J,* 22(13), pp. 3267-78.
Gallien, S., Duriez, E. and Domon, B. (2011) 'Selected reaction monitoring applied to proteomics', *J Mass Spectrom,* 46(3), pp. 298-312.
Gao, P., Liu, S., Yoshida, R., Shi, C. Y., Yoshimachi, S., Sakata, N., Goto, K., Kimura, T., Shirakawa, R., Nakayama, H., Sakata, J., Kawashiri, S., Kato, K., Wang, X. Y. and Horiuchi, H. (2019) 'Ral GTPase Activation by Downregulation of RalGAP Enhances Oral Squamous Cell Carcinoma Progression', *J Dent Res,* 98(9), pp. 1011-1019.
Gao, Y., Dickerson, J. B., Guo, F., Zheng, J. and Zheng, Y. (2004) 'Rational design and characterization of a Rac GTPase-specific small molecule inhibitor', *Proc Natl Acad Sci U S A,* 101(20), pp. 7618-23.
Gauto, D. F., Estrozi, L. F., Schwieters, C. D., Effantin, G., Macek, P., Sounier, R., Sivertsen, A. C., Schmidt, E., Kerfah, R., Mas, G., Colletier, J. P., Güntert, P., Favier, A., Schoehn, G., Schanda, P. and Boisbouvier, J. (2019) 'Integrated NMR and cryo-EM atomic-resolution structure determination of a half-megadalton enzyme complex', *Nat Commun,* 10(1), pp. 2697.
Geenen, J. J. J. and Schellens, J. H. M. (2017) 'Molecular Pathways: Targeting the Protein Kinase Wee1 in Cancer', *Clin Cancer Res,* 23(16), pp. 4540-4544.
Geiger, T. and Clarke, S. (1987) 'Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation', *J Biol Chem,* 262(2), pp. 785-94.
Gentry, L. R., Martin, T. D., Reiner, D. J. and Der, C. J. (2014) 'Ral small GTPase signaling and oncogenesis: More than just 15minutes of fame', *Biochim Biophys Acta,* 1843(12), pp. 2976-2988.
Gentry, L. R., Nishimura, A., Cox, A. D., Martin, T. D., Tsygankov, D., Nishida, M., Elston, T. C. and Der, C. J. (2015) 'Divergent roles of CAAX motif-signaled posttranslational modifications in the regulation and subcellular localization of Ral GTPases', *J Biol Chem,* 290(37), pp. 22851-61.
Gerecke, C., Mascher, C., Gottschalk, U., Kleuser, B. and Scholtka, B. (2013) 'Ultrasensitive detection of unknown colon cancer-initiating mutations using the example of the Adenomatous polyposis coli gene', *Cancer Prev Res (Phila),* 6(9), pp. 898-907.
Ghesquière, B. and Gevaert, K. (2014) 'Proteomics methods to study methionine oxidation', *Mass Spectrom Rev,* 33(2), pp. 147-56.

Ginn, K. F., Fangman, B., Terai, K., Wise, A., Ziazadeh, D., Shah, K., Gartrell, R., Ricke, B., Kimura, K., Mathur, S., Borrego-Diaz, E. and Farassati, F. (2016) 'RalA is overactivated in medulloblastoma', *J Neurooncol,* 130(1), pp. 99-110.
Goi, T., Shipitsin, M., Lu, Z., Foster, D. A., Klinz, S. G. and Feig, L. A. (2000) 'An EGF receptor/Ral-GTPase signaling cascade regulates c-Src activity and substrate specificity', *EMBO J,* 19(4), pp. 623-30.
Goody, R. S., Frech, M. and Wittinghofer, A. (1991) 'Affinity of guanine nucleotide binding proteins for their ligands: facts and artefacts', *Trends Biochem Sci,* 16(9), pp. 327-8.
Grabocka, E. and Bar-Sagi, D. (2016) 'Mutant KRAS Enhances Tumor Cell Fitness by Upregulating Stress Granules', *Cell,* 167(7), pp. 1803-1813.e12.
Guin, S., Ru, Y., Wynes, M. W., Mishra, R., Lu, X., Owens, C., Barn, A. E., Vasu, V. T., Hirsch, F. R., Kern, J. A. and Theodorescu, D. (2013) 'Contributions of KRAS and RAL in non-small-cell lung cancer growth and progression', *J Thorac Oncol,* 8(12), pp. 1492-501.
Gupta, R. M. and Musunuru, K. (2014) 'Expanding the genetic editing tool kit: ZFNs, TALENs, and CRISPR-Cas9', *J Clin Invest,* 124(10), pp. 4154-61.
Győrffy, B., Stelniec-Klotz, I., Sigler, C., Kasack, K., Redmer, T., Qian, Y. and Schäfer, R. (2015) 'Effects of RAL signal transduction in KRAS- and BRAF-mutated cells and prognostic potential of the RAL signature in colorectal cancer', *Oncotarget,* 6(15), pp. 13334-46.
Haber, J. E., Borts, R. H., Connolly, B., Lichten, M., Rudin, N. and White, C. I. (1988) 'Physical monitoring of meiotic and mitotic recombination in yeast', *Prog Nucleic Acid Res Mol Biol,* 35, pp. 209-59.
Hallin, J., Engstrom, L. D., Hargis, L., Calinisan, A., Aranda, R., Briere, D. M., Sudhakar, N., Bowcut, V., Baer, B. R., Ballard, J. A., Burkard, M. R., Fell, J. B., Fischer, J. P., Vigers, G. P., Xue, Y., Gatto, S., Fernandez-Banet, J., Pavlicek, A., Velastagui, K., Chao, R. C., Barton, J., Pierobon, M., Baldelli, E., Patricoin, E. F., Cassidy, D. P., Marx, M. A., Rybkin, I. I., Johnson, M. L., Ou, S. I., Lito, P., Papadopoulos, K. P., Jänne, P. A., Olson, P. and Christensen, J. G. (2020) 'The KRAS', *Cancer Discov,* 10(1), pp. 54-71.
Halvey, P. J., Ferrone, C. R. and Liebler, D. C. (2012) 'GeLC-MRM quantitation of mutant KRAS oncoprotein in complex biological samples', *J Proteome Res,* 11(7), pp. 3908-13.
Hamad, N. M., Elconin, J. H., Karnoub, A. E., Bai, W., Rich, J. N., Abraham, R. T., Der, C. J. and Counter, C. M. (2002) 'Distinct requirements for Ras oncogenesis in human versus mouse cells', *Genes Dev,* 16(16), pp. 2045-57.
Hamada, M., Miki, T., Iwai, S., Shimizu, H. and Yura, Y. (2011) 'Involvement of RhoA and RalB in geranylgeranyltransferase I inhibitor-mediated inhibition of proliferation and migration of human oral squamous cell carcinoma cells', *Cancer Chemother Pharmacol,* 68(3), pp. 559-69.
He, B. and Guo, W. (2009) 'The exocyst complex in polarized exocytosis', *Curr Opin Cell Biol,* 21(4), pp. 537-42.
Hennig, A., Markwart, R., Wolff, K., Schubert, K., Cui, Y., Prior, I. A., Esparza-Franco, M. A., Ladds, G. and Rubio, I. (2016) 'Feedback activation of neurofibromin terminates growth factor-induced Ras activation', *Cell Commun Signal,* 14, pp. 5.
Henry, D. O., Moskalenko, S. A., Kaur, K. J., Fu, M., Pestell, R. G., Camonis, J. H. and White, M. A. (2000) 'Ral GTPases contribute to regulation of cyclin D1 through activation of NF-kappaB', *Mol Cell Biol,* 20(21), pp. 8084-92.

Hofer, F., Fields, S., Schneider, C. and Martin, G. S. (1994) 'Activated Ras interacts with the Ral guanine nucleotide dissociation stimulator', *Proc Natl Acad Sci U S A,* 91(23), pp. 11089-93.
Hood, F. E., Klinger, B., Newlaczyl, A. U., Sieber, A., Dorel, M., Oliver, S. P., Coulson, J. M., Blüthgen, N. and Prior, I. A. (2019) 'Isoform-specific Ras signaling is growth factor dependent', *Mol Biol Cell,* 30(9), pp. 1108-1117.
Howe, J. R., Skryabin, B. V., Belcher, S. M., Zerillo, C. A. and Schmauss, C. (1995) 'The responsiveness of a tetracycline-sensitive expression system differs in different cell lines', *J Biol Chem,* 270(23), pp. 14168-74.
Hurd, C. A., Brear, P., Revell, J., Ross, S., Mott, H. R. and Owen, D. (2020) 'Affinity maturation of the RLIP76 Ral binding domain to inform the design of stapled peptides targeting the Ral GTPases', *J Biol Chem,* 296, pp. 100101.
Ieong, C., Ma, J. and Lai, W. (2019) 'RALBP1 regulates oral cancer cells via Akt and is a novel target of miR-148a-3p and miR-148b-3p', *J Oral Pathol Med,* 48(10), pp. 919-928.
Ihle, N. T., Byers, L. A., Kim, E. S., Saintigny, P., Lee, J. J., Blumenschein, G. R., Tsao, A., Liu, S., Larsen, J. E., Wang, J., Diao, L., Coombes, K. R., Chen, L., Zhang, S., Abdelmelek, M. F., Tang, X., Papadimitrakopoulou, V., Minna, J. D., Lippman, S. M., Hong, W. K., Herbst, R. S., Wistuba, I. I., Heymach, J. V. and Powis, G. (2012) 'Effect of KRAS oncogene substitutions on protein behavior: implications for signaling and clinical outcome', *J Natl Cancer Inst,* 104(3), pp. 228-39.
Ikeda, M., Ishida, O., Hinoi, T., Kishida, S. and Kikuchi, A. (1998) 'Identification and characterization of a novel protein interacting with Ral-binding protein 1, a putative effector protein of Ral', *J Biol Chem,* 273(2), pp. 814-21.
Isomura, M., Okui, K., Fujiwara, T., Shin, S. and Nakamura, Y. (1996) 'Isolation and mapping of RAB2L, a human cDNA that encodes a protein homologous to RalGDS', *Cytogenet Cell Genet,* 74(4), pp. 263-5.
Jeffrey, K. L., Brummer, T., Rolph, M. S., Liu, S. M., Callejas, N. A., Grumont, R. J., Gillieron, C., Mackay, F., Grey, S., Camps, M., Rommel, C., Gerondakis, S. D. and Mackay, C. R. (2006) 'Positive regulation of immune cell function and inflammatory responses by phosphatase PAC-1', *Nat Immunol,* 7(3), pp. 274-83.
Jiang, W., Bikard, D., Cox, D., Zhang, F. and Marraffini, L. A. (2013) 'RNA-guided editing of bacterial genomes using CRISPR-Cas systems', *Nat Biotechnol,* 31(3), pp. 233-9.
Jin, R., Junutula, J. R., Matern, H. T., Ervin, K. E., Scheller, R. H. and Brunger, A. T. (2005) 'Exo84 and Sec5 are competitive regulatory Sec6/8 effectors to the RalA GTPase', *EMBO J,* 24(12), pp. 2064-74.
Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012) 'A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity', *Science,* 337(6096), pp. 816-21.
Jinek, M., East, A., Cheng, A., Lin, S., Ma, E. and Doudna, J. (2013) 'RNA-programmed genome editing in human cells', *Elife,* 2, pp. e00471.
Johansson, J., Naszai, M., Hodder, M. C., Pickering, K. A., Miller, B. W., Ridgway, R. A., Yu, Y., Peschard, P., Brachmann, S., Campbell, A. D., Cordero, J. B. and Sansom, O. J. (2019) 'RAL GTPases Drive Intestinal Stem Cell Function and Regeneration through Internalization of WNT Signalosomes', *Cell Stem Cell,* 24(4), pp. 592-607.e7.

Jones, S., Zhang, X., Parsons, D. W., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Kamiyama, H., Jimeno, A., Hong, S. M., Fu, B., Lin, M. T., Calhoun, E. S., Kamiyama, M., Walter, K., Nikolskaya, T., Nikolsky, Y., Hartigan, J., Smith, D. R., Hidalgo, M., Leach, S. D., Klein, A. P., Jaffee, E. M., Goggins, M., Maitra, A., Iacobuzio-Donahue, C., Eshleman, J. R., Kern, S. E., Hruban, R. H., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E. and Kinzler, K. W. (2008) 'Core signaling pathways in human pancreatic cancers revealed by global genomic analyses', *Science,* 321(5897), pp. 1801-6.

Joung, J. K. and Sander, J. D. (2013) 'TALENs: a widely applicable technology for targeted genome editing', *Nat Rev Mol Cell Biol,* 14(1), pp. 49-55.

Jullien-Flores, V., Dorseuil, O., Romero, F., Letourneur, F., Saragosti, S., Berger, R., Tavitian, A., Gacon, G. and Camonis, J. H. (1995) 'Bridging Ral GTPase to Rho pathways. RLIP76, a Ral effector with CDC42/Rac GTPase-activating protein activity', *J Biol Chem,* 270(38), pp. 22473-7.

Jullien-Flores, V., Mahé, Y., Mirey, G., Leprince, C., Meunier-Bisceuil, B., Sorkin, A. and Camonis, J. H. (2000) 'RLIP76, an effector of the GTPase Ral, interacts with the AP2 complex: involvement of the Ral pathway in receptor endocytosis', *J Cell Sci,* 113 (Pt 16), pp. 2837-44.

Karunanithi, S., Xiong, T., Uhm, M., Leto, D., Sun, J., Chen, X. W. and Saltiel, A. R. (2014) 'A Rab10:RalA G protein cascade regulates insulin-stimulated glucose uptake in adipocytes', *Mol Biol Cell,* 25(19), pp. 3059-69.

Kashatus, D. F., Lim, K. H., Brady, D. C., Pershing, N. L., Cox, A. D. and Counter, C. M. (2011) 'RALA and RALBP1 regulate mitochondrial fission at mitosis', *Nat Cell Biol,* 13(9), pp. 1108-15.

Kidd, A. R., Snider, J. L., Martin, T. D., Graboski, S. F., Der, C. J. and Cox, A. D. (2010) 'Ras-related small GTPases RalA and RalB regulate cellular survival after ionizing radiation', *Int J Radiat Oncol Biol Phys,* 78(1), pp. 205-12.

Kikuchi, A., Demo, S. D., Ye, Z. H., Chen, Y. W. and Williams, L. T. (1994) 'ralGDS family members interact with the effector loop of ras p21', *Mol Cell Biol,* 14(11), pp. 7483-91.

Kim, H. and Kim, J. S. (2014) 'A guide to genome engineering with programmable nucleases', *Nat Rev Genet,* 15(5), pp. 321-34.

Kim, J. H., Lee, S. D., Han, J. M., Lee, T. G., Kim, Y., Park, J. B., Lambeth, J. D., Suh, P. G. and Ryu, S. H. (1998) 'Activation of phospholipase D1 by direct interaction with ADP-ribosylation factor 1 and RalA', *FEBS Lett,* 430(3), pp. 231-5.

Kim, K., Kuo, T., Cai, J., Shuja, S. and Murnane, M. J. (1997) 'N-ras protein: frequent quantitative and qualitative changes occur in human colorectal carcinomas', *Int J Cancer,* 71(5), pp. 767-75.

Kim, Y. G., Cha, J. and Chandrasegaran, S. (1996) 'Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain', *Proc Natl Acad Sci U S A,* 93(3), pp. 1156-60.

Kinsella, B. T., Erdman, R. A. and Maltese, W. A. (1991) 'Carboxyl-terminal isoprenylation of ras-related GTP-binding proteins encoded by rac1, rac2, and ralA', *J Biol Chem,* 266(15), pp. 9786-94.

Kohsaka, S., Hinohara, K., Wang, L., Nishimura, T., Urushido, M., Yachi, K., Tsuda, M., Tanino, M., Kimura, T., Nishihara, H., Gotoh, N. and Tanaka, S. (2014) 'Epiregulin enhances tumorigenicity by activating the ERK/MAPK pathway in glioblastoma', *Neuro Oncol,* 16(7), pp. 960-70.

Koike-Yusa, H., Li, Y., Tan, E. P., Velasco-Herrera, M. e. C. and Yusa, K. (2014) 'Genome-wide recessive genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library', *Nat Biotechnol,* 32(3), pp. 267-73.
Kou, C. J. and Kandpal, R. P. (2018) 'Differential Expression Patterns of Eph Receptors and Ephrin Ligands in Human Cancers', *Biomed Res Int,* 2018, pp. 7390104.
Krasinskas, A. M., Chiosea, S. I., Pal, T. and Dacic, S. (2014) 'KRAS mutational analysis and immunohistochemical studies can help distinguish pancreatic metastases from primary lung adenocarcinomas', *Mod Pathol,* 27(2), pp. 262-70.
Krasinskas, A. M., Moser, A. J., Saka, B., Adsay, N. V. and Chiosea, S. I. (2013) 'KRAS mutant allele-specific imbalance is associated with worse prognosis in pancreatic cancer and progression to undifferentiated carcinoma of the pancreas', *Mod Pathol,* 26(10), pp. 1346-54.
Lake, D., Corrêa, S. A. and Müller, J. (2016) 'Negative feedback regulation of the ERK1/2 MAPK pathway', *Cell Mol Life Sci,* 73(23), pp. 4397-4413.
Lange, V., Picotti, P., Domon, B. and Aebersold, R. (2008) 'Selected reaction monitoring for quantitative proteomics: a tutorial', *Mol Syst Biol,* 4, pp. 222.
Lee, C. S., Lee, L. C., Yuan, T. L., Chakka, S., Fellmann, C., Lowe, S. W., Caplen, N. J., McCormick, F. and Luo, J. (2019) 'MAP kinase and autophagy pathways cooperate to maintain RAS mutant cancer cell survival', *Proc Natl Acad Sci U S A,* 116(10), pp. 4508-4517.
Lee, K. J., Hoe, H. S. and Pak, D. T. (2011) 'Plk2 Raps up Ras to subdue synapses', *Small GTPases,* 2(3), pp. 162-166.
Lee, K. J., Lee, Y., Rozeboom, A., Lee, J. Y., Udagawa, N., Hoe, H. S. and Pak, D. T. (2011) 'Requirement for Plk2 in orchestrated ras and rap signaling, homeostatic structural plasticity, and memory', *Neuron,* 69(5), pp. 957-73.
Lim, K. H., Baines, A. T., Fiordalisi, J. J., Shipitsin, M., Feig, L. A., Cox, A. D., Der, C. J. and Counter, C. M. (2005) 'Activation of RalA is critical for Ras-induced tumorigenesis of human cells', *Cancer Cell,* 7(6), pp. 533-45.
Lim, K. H., Brady, D. C., Kashatus, D. F., Ancrile, B. B., Der, C. J., Cox, A. D. and Counter, C. M. (2010) 'Aurora-A phosphorylates, activates, and relocalizes the small GTPase RalA', *Mol Cell Biol,* 30(2), pp. 508-23.
Lim, K. H., O'Hayer, K., Adam, S. J., Kendall, S. D., Campbell, P. M., Der, C. J. and Counter, C. M. (2006) 'Divergent roles for RalA and RalB in malignant growth of human pancreatic carcinoma cells', *Curr Biol,* 16(24), pp. 2385-94.
Liu, C., Zhang, L., Liu, H. and Cheng, K. (2017) 'Delivery strategies of the CRISPR-Cas9 gene-editing system for therapeutic applications', *J Control Release,* 266, pp. 17-26.
Liu, J., Srinivasan, S., Li, C. Y., Ho, I. L., Rose, J., Shaheen, M., Wang, G., Yao, W., Deem, A., Bristow, C., Hart, T. and Draetta, G. (2019) 'Pooled library screening with multiplexed Cpf1 library', *Nat Commun,* 10(1), pp. 3144.
Ljubicic, S., Bezzi, P., Vitale, N. and Regazzi, R. (2009) 'The GTPase RalA regulates different steps of the secretory process in pancreatic beta-cells', *PLoS One,* 4(11), pp. e7770.
Love, M. I., Huber, W. and Anders, S. (2014) 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol,* 15(12), pp. 550.

Lowenthal, M. S., Liang, Y., Phinney, K. W. and Stein, S. E. (2014) 'Quantitative bottom-up proteomics depends on digestion conditions', *Anal Chem,* 86(1), pp. 551-8.
Lu, A., Wawro, P., Morgens, D. W., Portela, F., Bassik, M. C. and Pfeffer, S. R. (2018) 'Genome-wide interrogation of extracellular vesicle biology using barcoded miRNAs', *Elife,* 7.
Luo, J. Q., Liu, X., Frankel, P., Rotunda, T., Ramos, M., Flom, J., Jiang, H., Feig, L. A., Morris, A. J., Kahn, R. A. and Foster, D. A. (1998) 'Functional association between Arf and RalA in active phospholipase D complex', *Proc Natl Acad Sci U S A,* 95(7), pp. 3632-7.
MacLeod, G., Bozek, D. A., Rajakulendran, N., Monteiro, V., Ahmadi, M., Steinhart, Z., Kushida, M. M., Yu, H., Coutinho, F. J., Cavalli, F. M. G., Restall, I., Hao, X., Hart, T., Luchman, H. A., Weiss, S., Dirks, P. B. and Angers, S. (2019) 'Genome-Wide CRISPR-Cas9 Screens Expose Genetic Vulnerabilities and Mechanisms of Temozolomide Sensitivity in Glioblastoma Stem Cells', *Cell Rep,* 27(3), pp. 971-986.e9.
Macrae, M., Neve, R. M., Rodriguez-Viciana, P., Haqq, C., Yeh, J., Chen, C., Gray, J. W. and McCormick, F. (2005) 'A conditional feedback loop regulates Ras activity through EphA2', *Cancer Cell,* 8(2), pp. 111-8.
Mageean, C. J., Griffiths, J. R., Smith, D. L., Clague, M. J. and Prior, I. A. (2015) 'Absolute Quantification of Endogenous Ras Isoform Abundance', *PLoS One,* 10(11), pp. e0142674.
Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E. and Church, G. M. (2013) 'RNA-guided human genome engineering via Cas9', *Science,* 339(6121), pp. 823-6.
Mallick, P., Schirle, M., Chen, S. S., Flory, M. R., Lee, H., Martin, D., Ranish, J., Raught, B., Schmitt, R., Werner, T., Kuster, B. and Aebersold, R. (2007) 'Computational prediction of proteotypic peptides for quantitative proteomics', *Nat Biotechnol,* 25(1), pp. 125-31.
Mansour, S. J., Matten, W. T., Hermann, A. S., Candia, J. M., Rong, S., Fukasawa, K., Vande Woude, G. F. and Ahn, N. G. (1994) 'Transformation of mammalian cells by constitutively active MAP kinase kinase', *Science,* 265(5174), pp. 966-70.
Martin, T. D., Chen, X. W., Kaplan, R. E., Saltiel, A. R., Walker, C. L., Reiner, D. J. and Der, C. J. (2014) 'Ral and Rheb GTPase activating proteins integrate mTOR and GTPase signaling in aging, autophagy, and tumor cell invasion', *Mol Cell,* 53(2), pp. 209-20.
Martin, T. D., Mitin, N., Cox, A. D., Yeh, J. J. and Der, C. J. (2012) 'Phosphorylation by protein kinase Cα regulates RalB small GTPase protein activation, subcellular localization, and effector utilization', *J Biol Chem,* 287(18), pp. 14827-36.
Martin, T. D., Samuel, J. C., Routh, E. D., Der, C. J. and Yeh, J. J. (2011) 'Activation and involvement of Ral GTPases in colorectal cancer', *Cancer Res,* 71(1), pp. 206-15.
Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., Linder, B., Pickering, B. F., Vasseur, J. J., Chen, Q., Gross, S. S., Elemento, O., Debart, F., Kiledjian, M. and Jaffrey, S. R. (2017) 'Reversible methylation of m', *Nature,* 541(7637), pp. 371-375.
Maurer, T., Garrenton, L. S., Oh, A., Pitts, K., Anderson, D. J., Skelton, N. J., Fauber, B. P., Pan, B., Malek, S., Stokoe, D., Ludlam, M. J., Bowman, K. K.,

Wu, J., Giannetti, A. M., Starovasnik, M. A., Mellman, I., Jackson, P. K., Rudolph, J., Wang, W. and Fang, G. (2012) 'Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange activity', *Proc Natl Acad Sci U S A,* 109(14), pp. 5299-304.
Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D. and Rebar, E. J. (2011) 'A TALE nuclease architecture for efficient genome editing', *Nat Biotechnol,* 29(2), pp. 143-8.
Mishra, P. J., Ha, L., Rieker, J., Sviderskaya, E. V., Bennett, D. C., Oberst, M. D., Kelly, K. and Merlino, G. (2010) 'Dissection of RAS downstream pathways in melanomagenesis: a role for Ral in transformation', *Oncogene,* 29(16), pp. 2449-56.
Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. and Soria, E. (2005) 'Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements', *J Mol Evol,* 60(2), pp. 174-82.
Morinaka, K., Koyama, S., Nakashima, S., Hinoi, T., Okawa, K., Iwamatsu, A. and Kikuchi, A. (1999) 'Epsin binds to the EH domain of POB1 and regulates receptor-mediated endocytosis', *Oncogene,* 18(43), pp. 5915-22.
Moriya, H. (2015) 'Quantitative nature of overexpression experiments', *Mol Biol Cell,* 26(22), pp. 3932-9.
Moskalenko, S., Henry, D. O., Rosse, C., Mirey, G., Camonis, J. H. and White, M. A. (2002) 'The exocyst is a Ral effector complex', *Nat Cell Biol,* 4(1), pp. 66-72.
Moskalenko, S., Tong, C., Rosse, C., Mirey, G., Formstecher, E., Daviet, L., Camonis, J. and White, M. A. (2003) 'Ral GTPases regulate exocyst assembly through dual subunit interactions', *J Biol Chem,* 278(51), pp. 51743-8.
Murovec, J., Pirc, Ž. and Yang, B. (2017) 'New variants of CRISPR RNA-guided genome editing enzymes', *Plant Biotechnol J,* 15(8), pp. 917-926.
Nakashima, S., Morinaka, K., Koyama, S., Ikeda, M., Kishida, M., Okawa, K., Iwamatsu, A., Kishida, S. and Kikuchi, A. (1999) 'Small G protein Ral and its downstream molecules regulate endocytosis of EGF and insulin receptors', *EMBO J,* 18(13), pp. 3629-42.
Neel, N. F., Rossman, K. L., Martin, T. D., Hayes, T. K., Yeh, J. J. and Der, C. J. (2012) 'The RalB small GTPase mediates formation of invadopodia through a GTPase-activating protein-independent function of the RalBP1/RLIP76 effector', *Mol Cell Biol,* 32(8), pp. 1374-86.
Neyraud, V., Aushev, V. N., Hatzoglou, A., Meunier, B., Cascone, I. and Camonis, J. (2012) 'RalA and RalB proteins are ubiquitinated GTPases, and ubiquitinated RalA increases lipid raft exposure at the plasma membrane', *J Biol Chem,* 287(35), pp. 29397-405.
Nicely, N. I., Kosak, J., de Serrano, V. and Mattos, C. (2004) 'Crystal structures of Ral-GppNHp and Ral-GDP reveal two binding sites that are also present in Ras and Rap', *Structure,* 12(11), pp. 2025-36.
Ohta, Y., Suzuki, N., Nakamura, S., Hartwig, J. H. and Stossel, T. P. (1999) 'The small GTPase RalA targets filamin to induce filopodia', *Proc Natl Acad Sci U S A,* 96(5), pp. 2122-8.
Oldham, S. M., Clark, G. J., Gangarosa, L. M., Coffey, R. J. and Der, C. J. (1996) 'Activation of the Raf-1/MAP kinase cascade is not sufficient for Ras

transformation of RIE-1 epithelial cells', *Proc Natl Acad Sci U S A,* 93(14), pp. 6924-8.
Olsen, J. V. and Mann, M. (2004) 'Improved peptide identification in proteomics by two consecutive stages of mass spectrometric fragmentation', *Proc Natl Acad Sci U S A,* 101(37), pp. 13417-22.
Omerovic, J., Hammond, D. E., Clague, M. J. and Prior, I. A. (2008) 'Ras isoform abundance and signalling in human cancer cell lines', *Oncogene,* 27(19), pp. 2754-62.
Ortinski, P. I., O'Donovan, B., Dong, X. and Kantor, B. (2017) 'Integrase-Deficient Lentiviral Vector as an All-in-One Platform for Highly Efficient CRISPR/Cas9-Mediated Gene Editing', *Mol Ther Methods Clin Dev,* 5, pp. 153-164.
Osei-Sarfo, K., Martello, L., Ibrahim, S. and Pellicer, A. (2011) 'The human Rgr oncogene is overexpressed in T-cell malignancies and induces transformation by acting as a GEF for Ras and Ral', *Oncogene,* 30(34), pp. 3661-71.
Ostrem, J. M., Peters, U., Sos, M. L., Wells, J. A. and Shokat, K. M. (2013) 'K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions', *Nature,* 503(7477), pp. 548-51.
Oxford, G., Owens, C. R., Titus, B. J., Foreman, T. L., Herlevsen, M. C., Smith, S. C. and Theodorescu, D. (2005) 'RalA and RalB: antagonistic relatives in cancer cell migration', *Cancer Res,* 65(16), pp. 7111-20.
Paizs, B. and Suhai, S. (2001) 'Theoretical study of the main fragmentation pathways for protonated glycylglycine', *Rapid Commun Mass Spectrom,* 15(8), pp. 651-63.
Panganiban, R. A., Park, H. R., Sun, M., Shumyatcher, M., Himes, B. E. and Lu, Q. (2019) 'Genome-wide CRISPR screen identifies suppressors of endoplasmic reticulum stress-induced apoptosis', *Proc Natl Acad Sci U S A,* 116(27), pp. 13384-13393.
Papas, T. S., Watson, D. K., Sacchi, N., Fujiwara, S., Seth, A. K., Fisher, R. J., Bhat, N. K., Mavrothalassitis, G., Koizumi, S. and Jorcyk, C. L. (1990) 'ETS family of genes in leukemia and Down syndrome', *Am J Med Genet Suppl,* 7, pp. 251-61.
Paquet, D., Kwart, D., Chen, A., Sproul, A., Jacob, S., Teo, S., Olsen, K. M., Gregg, A., Noggle, S. and Tessier-Lavigne, M. (2016) 'Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9', *Nature,* 533(7601), pp. 125-9.
Park, S. H. and Weinberg, R. A. (1995) 'A putative effector of Ral has homology to Rho/Rac GTPase activating proteins', *Oncogene,* 11(11), pp. 2349-55.
Personnic, N., Lakisic, G., Gouin, E., Rousseau, A., Gautreau, A., Cossart, P. and Bierne, H. (2014) 'A role for Ral GTPase-activating protein subunit β in mitotic regulation', *FEBS J,* 281(13), pp. 2977-89.
Peschard, P., McCarthy, A., Leblanc-Dominguez, V., Yeo, M., Guichard, S., Stamp, G. and Marshall, C. J. (2012) 'Genetic deletion of RALA and RALB small GTPases reveals redundant functions in development and tumorigenesis', *Curr Biol,* 22(21), pp. 2063-8.
Petris, G., Casini, A., Montagna, C., Lorenzin, F., Prandi, D., Romanel, A., Zasso, J., Conti, L., Demichelis, F. and Cereseto, A. (2017) 'Hit and go CAS9 delivered through a lentiviral based self-limiting circuit', *Nat Commun,* 8, pp. 15334.

Piszkiewicz, D., Landon, M. and Smith, E. L. (1970) 'Anomalous cleavage of aspartyl-proline peptide bonds during amino acid sequence determinations', *Biochem Biophys Res Commun,* 40(5), pp. 1173-8.
Poirier, J. T. (2017) 'CRISPR Libraries and Screening', *Prog Mol Biol Transl Sci,* 152, pp. 69-82.
Pollock, S. R., Schinlever, A. R., Rohani, A., Kashatus, J. A. and Kashatus, D. F. (2019) 'RalA and RalB relocalization to depolarized mitochondria depends on clathrin-mediated endocytosis and facilitates TBK1 activation', *PLoS One,* 14(4), pp. e0214764.
Prior, I. A., Lewis, P. D. and Mattos, C. (2012) 'A comprehensive survey of Ras mutations in cancer', *Cancer Res,* 72(10), pp. 2457-67.
Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y. and Zhang, F. (2013a) 'Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity', *Cell,* 154(6), pp. 1380-9.
Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A. and Zhang, F. (2013b) 'Genome engineering using the CRISPR-Cas9 system', *Nat Protoc,* 8(11), pp. 2281-2308.
Rebhun, J. F., Chen, H. and Quilliam, L. A. (2000) 'Identification and characterization of a new family of guanine nucleotide exchange factors for the ras-related GTPase Ral', *J Biol Chem,* 275(18), pp. 13406-10.
Reddy, A., Zhang, J., Davis, N. S., Moffitt, A. B., Love, C. L., Waldrop, A., Leppa, S., Pasanen, A., Meriranta, L., Karjalainen-Lindsberg, M. L., Nørgaard, P., Pedersen, M., Gang, A. O., Høgdall, E., Heavican, T. B., Lone, W., Iqbal, J., Qin, Q., Li, G., Kim, S. Y., Healy, J., Richards, K. L., Fedoriw, Y., Bernal-Mizrachi, L., Koff, J. L., Staton, A. D., Flowers, C. R., Paltiel, O., Goldschmidt, N., Calaminici, M., Clear, A., Gribben, J., Nguyen, E., Czader, M. B., Ondrejka, S. L., Collie, A., Hsi, E. D., Tse, E., Au-Yeung, R. K. H., Kwong, Y. L., Srivastava, G., Choi, W. W. L., Evens, A. M., Pilichowska, M., Sengar, M., Reddy, N., Li, S., Chadburn, A., Gordon, L. I., Jaffe, E. S., Levy, S., Rempel, R., Tzeng, T., Happ, L. E., Dave, T., Rajagopalan, D., Datta, J., Dunson, D. B. and Dave, S. S. (2017) 'Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma', *Cell,* 171(2), pp. 481-494.e15.
Renaud, J. B., Boix, C., Charpentier, M., De Cian, A., Cochennec, J., Duvernois-Berthet, E., Perrouault, L., Tesson, L., Edouard, J., Thinard, R., Cherifi, Y., Menoret, S., Fontanière, S., de Crozé, N., Fraichard, A., Sohm, F., Anegon, I., Concordet, J. P. and Giovannangeli, C. (2016) 'Improved Genome Editing Efficiency and Flexibility Using Modified Oligonucleotides with TALEN and CRISPR-Cas9 Nucleases', *Cell Rep,* 14(9), pp. 2263-2272.
Renault, L., Guibert, B. and Cherfils, J. (2003) 'Structural snapshots of the mechanism and inhibition of a guanine nucleotide exchange factor', *Nature,* 426(6966), pp. 525-30.
Richardson, C. D., Ray, G. J., DeWitt, M. A., Curie, G. L. and Corn, J. E. (2016) 'Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA', *Nat Biotechnol,* 34(3), pp. 339-44.
Rodriguez-Viciana, P., Sabatier, C. and McCormick, F. (2004) 'Signaling specificity by Ras family GTPases is determined by the full spectrum of effectors they regulate', *Mol Cell Biol,* 24(11), pp. 4943-54.

Rogers, M. F., Shihab, H. A., Mort, M., Cooper, D. N., Gaunt, T. R. and Campbell, C. (2018) 'FATHMM-XF: accurate prediction of pathogenic point mutations via extended features', *Bioinformatics,* 34(3), pp. 511-513.
Rossé, C., L'Hoste, S., Offner, N., Picard, A. and Camonis, J. (2003) 'RLIP, an effector of the Ral GTPases, is a platform for Cdk1 to phosphorylate epsin during the switch off of endocytosis in mitosis', *J Biol Chem,* 278(33), pp. 30597-604.
Rouet, P., Smih, F. and Jasin, M. (1994) 'Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease', *Mol Cell Biol,* 14(12), pp. 8096-106.
Sablina, A. A., Chen, W., Arroyo, J. D., Corral, L., Hector, M., Bulmer, S. E., DeCaprio, J. A. and Hahn, W. C. (2007) 'The tumor suppressor PP2A Abeta regulates the RalA GTPase', *Cell,* 129(5), pp. 969-82.
Saito, R., Shirakawa, R., Nishiyama, H., Kobayashi, T., Kawato, M., Kanno, T., Nishizawa, K., Matsui, Y., Ohbayashi, T., Horiguchi, M., Nakamura, T., Ikeda, T., Yamane, K., Nakayama, E., Nakamura, E., Toda, Y., Kimura, T., Kita, T., Ogawa, O. and Horiuchi, H. (2013) 'Downregulation of Ral GTPase-activating protein promotes tumor invasion and metastasis of bladder cancer', *Oncogene,* 32(7), pp. 894-902.
Schmidt, S., Diriong, S., Méry, J., Fabbrizio, E. and Debant, A. (2002) 'Identification of the first Rho-GEF inhibitor, TRIPalpha, which targets the RhoA-specific GEF domain of Trio', *FEBS Lett,* 523(1-3), pp. 35-42.
Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W. and Selbach, M. (2011) 'Global quantification of mammalian gene expression control', *Nature,* 473(7347), pp. 337-42.
Seeburg, P. H., Colby, W. W., Capon, D. J., Goeddel, D. V. and Levinson, A. D. (1984) 'Biological properties of human c-Ha-ras1 genes mutated at codon 12', *Nature,* 312(5989), pp. 71-5.
Serebriiskii, I. G., Connelly, C., Frampton, G., Newberg, J., Cooke, M., Miller, V., Ali, S., Ross, J. S., Handorf, E., Arora, S., Lieu, C., Golemis, E. A. and Meyer, J. E. (2019) 'Comprehensive characterization of RAS mutations in colon and rectal cancers in old and young patients', *Nat Commun,* 10(1), pp. 3722.
Shalem, O., Sanjana, N. E., Hartenian, E., Shi, X., Scott, D. A., Mikkelson, T., Heckl, D., Ebert, B. L., Root, D. E., Doench, J. G. and Zhang, F. (2014) 'Genome-scale CRISPR-Cas9 knockout screening in human cells', *Science,* 343(6166), pp. 84-87.
Shao, H. and Andres, D. A. (2000) 'A novel RalGEF-like protein, RGL3, as a candidate effector for rit and Ras', *J Biol Chem,* 275(35), pp. 26914-24.
Shao, S., Ren, C., Liu, Z., Bai, Y., Chen, Z., Wei, Z., Wang, X., Zhang, Z. and Xu, K. (2017) 'Enhancing CRISPR/Cas9-mediated homology-directed repair in mammalian cells by expressing Saccharomyces cerevisiae Rad52', *Int J Biochem Cell Biol,* 92, pp. 43-52.
Shi, C. S., Shenderov, K., Huang, N. N., Kabat, J., Abu-Asab, M., Fitzgerald, K. A., Sher, A. and Kehrl, J. H. (2012) 'Activation of autophagy by inflammatory signals limits IL-1β production by targeting ubiquitinated inflammasomes for destruction', *Nat Immunol,* 13(3), pp. 255-63.
Shima, F., Ijiri, Y., Muraoka, S., Liao, J., Ye, M., Araki, M., Matsumoto, K., Yamamoto, N., Sugimoto, T., Yoshikawa, Y., Kumasaka, T., Yamamoto, M.,

Tamura, A. and Kataoka, T. (2010) 'Structural basis for conformational dynamics of GTP-bound Ras protein', *J Biol Chem,* 285(29), pp. 22696-705.
Shima, F., Yoshikawa, Y., Ye, M., Araki, M., Matsumoto, S., Liao, J., Hu, L., Sugimoto, T., Ijiri, Y., Takeda, A., Nishiyama, Y., Sato, C., Muraoka, S., Tamura, A., Osoda, T., Tsuda, K., Miyakawa, T., Fukunishi, H., Shimada, J., Kumasaka, T., Yamamoto, M. and Kataoka, T. (2013) 'In silico discovery of small-molecule Ras inhibitors that display antitumor activity by blocking the Ras-effector interaction', *Proc Natl Acad Sci U S A,* 110(20), pp. 8182-7.
Shipitsin, M. and Feig, L. A. (2004) 'RalA but not RalB enhances polarized delivery of membrane proteins to the basolateral surface of epithelial cells', *Mol Cell Biol,* 24(13), pp. 5746-56.
Shirakawa, R., Fukai, S., Kawato, M., Higashi, T., Kondo, H., Ikeda, T., Nakayama, E., Okawa, K., Nureki, O., Kimura, T., Kita, T. and Horiuchi, H. (2009) 'Tuberous sclerosis tumor suppressor complex-like complexes act as GTPase-activating proteins for Ral GTPases', *J Biol Chem,* 284(32), pp. 21580-8.
Sigoillot, F. D. and King, R. W. (2011) 'Vigilance and validation: Keys to success in RNAi screening', *ACS Chem Biol,* 6(1), pp. 47-60.
Sim, S. E., Lee, H. R., Kim, J. I., Choi, S. L., Bakes, J., Jang, D. J., Lee, K., Han, K., Kim, E. and Kaang, B. K. (2013) 'Elevated RalA activity in the hippocampus of PI3Kγ knock-out mice lacking NMDAR-dependent long-term depression', *BMB Rep,* 46(2), pp. 103-6.
Simicek, M., Lievens, S., Laga, M., Guzenko, D., Aushev, V. N., Kalev, P., Baietti, M. F., Strelkov, S. V., Gevaert, K., Tavernier, J. and Sablina, A. A. (2013) 'The deubiquitylase USP33 discriminates between RALB functions in autophagy and innate immune response', *Nat Cell Biol,* 15(10), pp. 1220-30.
Singhal, S. S., Salgia, R., Singhal, S., Horne, D. and Awasthi, S. (2019) 'RLIP: An existential requirement for breast carcinogenesis', *Biochim Biophys Acta Rev Cancer,* 1871(2), pp. 281-288.
Singhal, S. S., Singhal, J., Yadav, S., Dwivedi, S., Boor, P. J., Awasthi, Y. C. and Awasthi, S. (2007) 'Regression of lung and colon cancer xenografts by depleting or inhibiting RLIP76 (Ral-binding protein 1)', *Cancer Res,* 67(9), pp. 4382-9.
Sloan, S. R., Newcomb, E. W. and Pellicer, A. (1990) 'Neutron radiation can activate K-ras via a point mutation in codon 146 and induces a different spectrum of ras mutations than does gamma radiation', *Mol Cell Biol,* 10(1), pp. 405-8.
Smith, S. C., Baras, A. S., Owens, C. R., Dancik, G. and Theodorescu, D. (2012) 'Transcriptional signatures of Ral GTPase are associated with aggressive clinicopathologic characteristics in human cancer', *Cancer Res,* 72(14), pp. 3480-91.
Smith, S. C., Oxford, G., Baras, A. S., Owens, C., Havaleshko, D., Brautigan, D. L., Safo, M. K. and Theodorescu, D. (2007) 'Expression of ral GTPases, their effectors, and activators in human bladder cancer', *Clin Cancer Res,* 13(13), pp. 3803-13.
Soh, J., Okumura, N., Lockwood, W. W., Yamamoto, H., Shigematsu, H., Zhang, W., Chari, R., Shames, D. S., Tang, X., MacAulay, C., Varella-Garcia, M., Vooder, T., Wistuba, I. I., Lam, S., Brekken, R., Toyooka, S., Minna, J. D., Lam, W. L. and Gazdar, A. F. (2009) 'Oncogene mutations, copy number gains

and mutant allele specific imbalance (MASI) frequently occur together in tumor cells', *PLoS One,* 4(10), pp. e7464.
Song, X., Hua, L., Xu, Y., Fang, Z., Wang, Y., Gao, J., Shi, Q., Zhou, X. and Yu, R. (2015) 'Involvement of RalB in the effect of geranylgeranyltransferase I on glioma cell migration and invasion', *Clin Transl Oncol,* 17(6), pp. 477-85.
Sowalsky, A. G., Alt-Holland, A., Shamis, Y., Garlick, J. A. and Feig, L. A. (2010) 'RalA suppresses early stages of Ras-induced squamous cell carcinoma progression', *Oncogene,* 29(1), pp. 45-55.
Spiczka, K. S. and Yeaman, C. (2008) 'Ral-regulated interaction between Sec5 and paxillin targets Exocyst to focal complexes during cell migration', *J Cell Sci,* 121(Pt 17), pp. 2880-91.
Spurrier, B., Ramalingam, S. and Nishizuka, S. (2008) 'Reverse-phase protein lysate microarrays for cell signaling analysis', *Nat Protoc,* 3(11), pp. 1796-808.
Sugihara, K., Asano, S., Tanaka, K., Iwamatsu, A., Okawa, K. and Ohta, Y. (2002) 'The exocyst complex binds the small GTPase RalA to mediate filopodia formation', *Nat Cell Biol,* 4(1), pp. 73-8.
Sun, Q., Burke, J. P., Phan, J., Burns, M. C., Olejniczak, E. T., Waterson, A. G., Lee, T., Rossanese, O. W. and Fesik, S. W. (2012) 'Discovery of small molecules that bind to K-Ras and inhibit Sos-mediated activation', *Angew Chem Int Ed Engl,* 51(25), pp. 6140-3.
Sung, H., Kanchi, K. L., Wang, X., Hill, K. S., Messina, J. L., Lee, J. H., Kim, Y., Dees, N. D., Ding, L., Teer, J. K., Yang, S., Sarnaik, A. A., Sondak, V. K., Mulé, J. J., Wilson, R. K., Weber, J. S. and Kim, M. (2016) 'Inactivation of RASA1 promotes melanoma tumorigenesis via R-Ras activation', *Oncotarget,* 7(17), pp. 23885-96.
Taylor, A. L. (1970) 'Current linkage map of Escherichia coli', *Bacteriol Rev,* 34(2), pp. 155-75.
Taylor, S. J. and Shalloway, D. (1996) 'Cell cycle-dependent activation of Ras', *Curr Biol,* 6(12), pp. 1621-7.
Tazat, K., Harsat, M., Goldshmid-Shagal, A., Ehrlich, M. and Henis, Y. I. (2013) 'Dual effects of Ral-activated pathways on p27 localization and TGF-β signaling', *Mol Biol Cell,* 24(11), pp. 1812-24.
Tetsu, O. and McCormick, F. (2017) 'ETS-targeted therapy: can it substitute for MEK inhibitors?', *Clin Transl Med,* 6(1), pp. 16.
Thomas, J. C., Cooper, J. M., Clayton, N. S., Wang, C., White, M. A., Abell, C., Owen, D. and Mott, H. R. (2016) 'Inhibition of Ral GTPases Using a Stapled Peptide Approach', *J Biol Chem,* 291(35), pp. 18310-25.
Tian, X., Rusanescu, G., Hou, W., Schaffhausen, B. and Feig, L. A. (2002) 'PDK1 mediates growth factor-induced Ral-GEF activation by a kinase-independent mechanism', *EMBO J,* 21(6), pp. 1327-38.
Toyama, B. H. and Hetzer, M. W. (2013) 'Protein homeostasis: live long, won't prosper', *Nat Rev Mol Cell Biol,* 14(1), pp. 55-61.
Traverse, S., Gomez, N., Paterson, H., Marshall, C. and Cohen, P. (1992) 'Sustained activation of the mitogen-activated protein (MAP) kinase cascade may be required for differentiation of PC12 cells. Comparison of the effects of nerve growth factor and epidermal growth factor', *Biochem J,* 288 (Pt 2), pp. 351-5.
Ullah, N., Mansha, M. and Casey, P. J. (2016) 'Protein Geranylgeranyltransferase Type 1 as a Target in Cancer', *Curr Cancer Drug Targets,* 16(7), pp. 563-71.

Urano, T., Emkey, R. and Feig, L. A. (1996) 'Ral-GTPases mediate a distinct downstream signaling pathway from Ras that facilitates cellular transformation', *EMBO J,* 15(4), pp. 810-6.
Urnov, F. D., Rebar, E. J., Holmes, M. C., Zhang, H. S. and Gregory, P. D. (2010) 'Genome editing with engineered zinc finger nucleases', *Nat Rev Genet,* 11(9), pp. 636-46.
van Dam, T. J., Bos, J. L. and Snel, B. (2011) 'Evolution of the Ras-like small GTPases and their regulators', *Small GTPases,* 2(1), pp. 4-16.
van den Berg, M. C., van Gogh, I. J., Smits, A. M., van Triest, M., Dansen, T. B., Visscher, M., Polderman, P. E., Vliem, M. J., Rehmann, H. and Burgering, B. M. (2016) 'The small GTPase RALA controls c-Jun N-terminal kinase-mediated FOXO activation by regulation of a JIP1 scaffold complex', *J Biol Chem,* 291(3), pp. 1200.
Vaughn, C. P., Zobell, S. D., Furtado, L. V., Baker, C. L. and Samowitz, W. S. (2011) 'Frequency of KRAS, BRAF, and NRAS mutations in colorectal cancer', *Genes Chromosomes Cancer,* 50(5), pp. 307-12.
Vigil, D., Martin, T. D., Williams, F., Yeh, J. J., Campbell, S. L. and Der, C. J. (2010) 'Aberrant overexpression of the Rgl2 Ral small GTPase-specific guanine nucleotide exchange factor promotes pancreatic cancer growth through Ral-dependent and Ral-independent mechanisms', *J Biol Chem,* 285(45), pp. 34729-40.
Wagner, M., Skorobogatko, Y., Pode-Shakked, B., Powell, C. M., Alhaddad, B., Seibt, A., Barel, O., Heimer, G., Hoffmann, C., Demmer, L. A., Perilla-Young, Y., Remke, M., Wieczorek, D., Navaratnarajah, T., Lichtner, P., Klee, D., Shamseldin, H. E., Al Mutairi, F., Mayatepek, E., Strom, T., Meitinger, T., Alkuraya, F. S., Anikster, Y., Saltiel, A. R. and Distelmaier, F. (2020) 'Bi-allelic Variants in RALGAPA1 Cause Profound Neurodevelopmental Disability, Muscular Hypotonia, Infantile Spasms, and Feeding Abnormalities', *Am J Hum Genet,* 106(2), pp. 246-255.
Wake, C. T., Gudewicz, T., Porter, T., White, A. and Wilson, J. H. (1984) 'How damaged is the biologically active subpopulation of transfected DNA?', *Mol Cell Biol,* 4(3), pp. 387-98.
Walensky, L. D. and Bird, G. H. (2014) 'Hydrocarbon-stapled peptides: principles, practice, and progress', *J Med Chem,* 57(15), pp. 6275-88.
Walkup, W. G., Sweredoski, M. J., Graham, R. L., Hess, S. and Kennedy, M. B. (2018) 'Phosphorylation of synaptic GTPase-activating protein (synGAP) by polo-like kinase (Plk2) alters the ratio of its GAP activity toward HRas, Rap1 and Rap2 GTPases', *Biochem Biophys Res Commun,* 503(3), pp. 1599-1604.
Wang, B., Li, K., Wang, A., Reiser, M., Saunders, T., Lockey, R. F. and Wang, J. W. (2015) 'Highly efficient CRISPR/HDR-mediated knock-in for mouse embryonic stem cells and zygotes', *Biotechniques,* 59(4), pp. 201-2, 204, 206-8.
Wang, H., Owens, C., Chandra, N., Conaway, M. R., Brautigan, D. L. and Theodorescu, D. (2010) 'Phosphorylation of RalB is important for bladder cancer cell growth and metastasis', *Cancer Res,* 70(21), pp. 8760-9.
Wang, H., Qi, M. and Cutler, A. J. (1993) 'A simple method of preparing plant samples for PCR', *Nucleic Acids Res,* 21(17), pp. 4153-4.
Wang, Q., Chaerkady, R., Wu, J., Hwang, H. J., Papadopoulos, N., Kopelovich, L., Maitra, A., Matthaei, H., Eshleman, J. R., Hruban, R. H.,

Kinzler, K. W., Pandey, A. and Vogelstein, B. (2011) 'Mutant proteins as cancer-specific biomarkers', *Proc Natl Acad Sci U S A,* 108(6), pp. 2444-9.
Wennerberg, K., Rossman, K. L. and Der, C. J. (2005) 'The Ras superfamily at a glance', *J Cell Sci,* 118(Pt 5), pp. 843-6.
Wolthuis, R. M., Bauer, B., van 't Veer, L. J., de Vries-Smits, A. M., Cool, R. H., Spaargaren, M., Wittinghofer, A., Burgering, B. M. and Bos, J. L. (1996) 'RalGDS-like factor (Rlf) is a novel Ras and Rap 1A-associating protein', *Oncogene,* 13(2), pp. 353-62.
Wolthuis, R. M., Franke, B., van Triest, M., Bauer, B., Cool, R. H., Camonis, J. H., Akkerman, J. W. and Bos, J. L. (1998a) 'Activation of the small GTPase Ral in platelets', *Mol Cell Biol,* 18(5), pp. 2486-91.
Wolthuis, R. M., Zwartkruis, F., Moen, T. C. and Bos, J. L. (1998b) 'Ras-dependent activation of the small GTPase Ral', *Curr Biol,* 8(8), pp. 471-4.
Wu, J. C., Chen, T. Y., Yu, C. T., Tsai, S. J., Hsu, J. M., Tang, M. J., Chou, C. K., Lin, W. J., Yuan, C. J. and Huang, C. Y. (2005) 'Identification of V23RalA-Ser194 as a critical mediator for Aurora-A-induced cellular motility and transformation by small pool expression screening', *J Biol Chem,* 280(10), pp. 9013-22.
Xu, L., Salloum, D., Medlin, P. S., Saqcena, M., Yellen, P., Perrella, B. and Foster, D. A. (2011) 'Phospholipase D mediates nutrient input to mammalian target of rapamycin complex 1 (mTORC1)', *J Biol Chem,* 286(29), pp. 25477-86.
Yamaguchi, A., Urano, T., Goi, T. and Feig, L. A. (1997) 'An Eps homology (EH) domain protein that binds to the Ral-GTPase target, RalBP1', *J Biol Chem,* 272(50), pp. 31230-4.
Yan, C., Liu, D., Li, L., Wempe, M. F., Guin, S., Khanna, M., Meier, J., Hoffman, B., Owens, C., Wysoczynski, C. L., Nitz, M. D., Knabe, W. E., Ahmed, M., Brautigan, D. L., Paschal, B. M., Schwartz, M. A., Jones, D. N., Ross, D., Meroueh, S. O. and Theodorescu, D. (2014) 'Discovery and characterization of small molecules that target the GTPase Ral', *Nature,* 515(7527), pp. 443-7.
Yan, C. and Theodorescu, D. (2018) 'RAL GTPases: Biology and Potential as Therapeutic Targets in Cancer', *Pharmacol Rev,* 70(1), pp. 1-11.
Yang, J., Song, Q., Cai, Y., Wang, P., Wang, M. and Zhang, D. (2015) 'RLIP76-dependent suppression of PI3K/AKT/Bcl-2 pathway by miR-101 induces apoptosis in prostate cancer', *Biochem Biophys Res Commun,* 463(4), pp. 900-6.
Yin, J., Pollock, C., Tracy, K., Chock, M., Martin, P., Oberst, M. and Kelly, K. (2007) 'Activation of the RalGEF/Ral pathway promotes prostate cancer metastasis to bone', *Mol Cell Biol,* 27(21), pp. 7538-50.
Yuan, T. L., Amzallag, A., Bagni, R., Yi, M., Afghani, S., Burgan, W., Fer, N., Strathern, L. A., Powell, K., Smith, B., Waters, A. M., Drubin, D., Thomson, T., Liao, R., Greninger, P., Stein, G. T., Murchie, E., Cortez, E., Egan, R. K., Procter, L., Bess, M., Cheng, K. T., Lee, C. S., Lee, L. C., Fellmann, C., Stephens, R., Luo, J., Lowe, S. W., Benes, C. H. and McCormick, F. (2018) 'Differential Effector Engagement by Oncogenic KRAS', *Cell Rep,* 22(7), pp. 1889-1902.
Zaccara, S., Ries, R. J. and Jaffrey, S. R. (2019) 'Reading, writing and erasing mRNA methylation', *Nat Rev Mol Cell Biol,* 20(10), pp. 608-624.

Zhang, W. and Liu, H. T. (2002) 'MAPK signal pathways in the regulation of cell proliferation in mammalian cells', *Cell Res,* 12(1), pp. 9-18.
Zhang, Y., Blattman, J. N., Kennedy, N. J., Duong, J., Nguyen, T., Wang, Y., Davis, R. J., Greenberg, P. D., Flavell, R. A. and Dong, C. (2004) 'Regulation of innate and adaptive immune responses by MAP kinase phosphatase 5', *Nature,* 430(7001), pp. 793-7.
Zhao, W., Jamshidiha, M., Lanyon-Hogg, T., Recchi, C., Cota, E. and Tate, E. W. (2017) 'Direct Targeting of the Ras GTPase Superfamily Through Structure- Based Design', *Curr Top Med Chem,* 17(1), pp. 16-29.
Zhu, T., Ling, L. and Lobie, P. E. (2002) 'Identification of a JAK2-independent pathway regulating growth hormone (GH)-stimulated p44/42 mitogen-activated protein kinase activity. GH activation of Ral and phospholipase D is Src-dependent', *J Biol Chem,* 277(47), pp. 45592-603.
Zhuang, X., Veltri, D. P. and Long, E. O. (2019) 'Genome-Wide CRISPR Screen Reveals Cancer Cell Resistance to NK Cells Induced by NK-Derived IFN-γ', *Front Immunol,* 10, pp. 2879.
Zipfel, P. A., Brady, D. C., Kashatus, D. F., Ancrile, B. D., Tyler, D. S. and Counter, C. M. (2010) 'Ral activation promotes melanomagenesis', *Oncogene,* 29(34), pp. 4859-64.
Šlechtová, T., Gilar, M., Kalíková, K. and Tesařová, E. (2015) 'Insight into Trypsin Miscleavage: Comparison of Kinetic Constants of Problematic Peptide Sequences', *Anal Chem,* 87(15), pp. 7636-43.